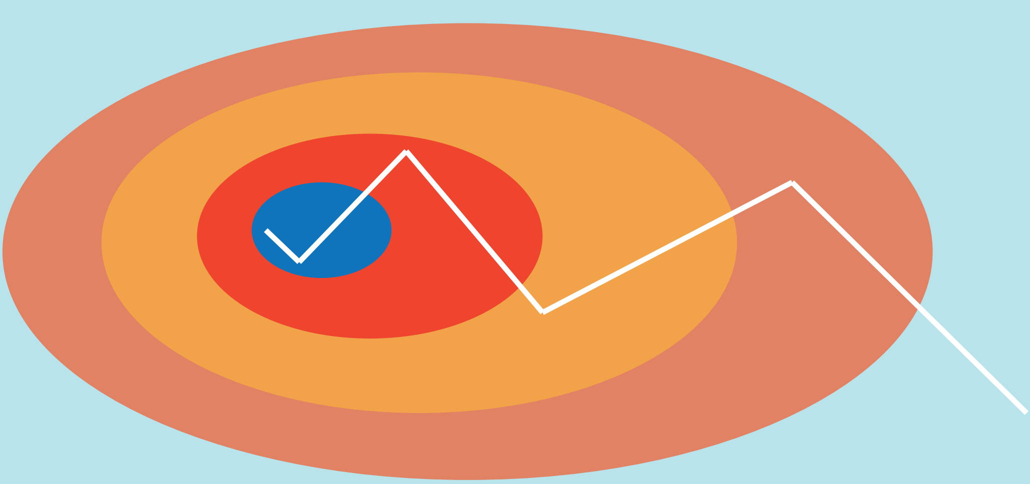
ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ
ՀԱՄԱԼՍԱՐԱՆ

ՌԱՖԻԿ ԽԱՉԱՏՐՅԱՆ

ՕՊՏԻՄԻԶԱՑԻԱՅԻ ՄԵԹՈԴՆԵՐ



ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

Ռ. Ա. ԽԱՉԱՏՐՅԱՆ

ՕՊՏԻՄԻԶԱՑԻԱՅԻ ՄԵԹՈԴՆԵՐ

ԵՐԵՎԱՆ
ԵՊՀ ՀՐԱՏԱՐԱԿՉՈՒԹՅՈՒՆ
2020

ՀՏԴ 519.85(075.8)
ԳՄԴ 22.18ց73
Խ 282

ՀՀ ԿԳՄՍ նախարարության կողմից
հաստատվել է որպես բուհական դասագիրք:

Հրատարակության է երաշխավորել
ԵՊՀ ինֆորմատիկայի և կիրառական մաթեմատիկայի
ֆակուլտետի գիտական խորհուրդը:

Խմբագիր՝ ֆիզ. մաթ. գիտ. թեկնածու, դոցենտ
Ս. Լ. Սահակյան (ԵՊՀ)

Խաչատրյան Ռ. Ա.
Խ 282 Օպտիմիզացիայի մեթոդներ: Դասագիրք/Ռ. Ա. Խաչատրյան.- Եր.: ԵՊՀ հրատ., 2020, 308 էջ:

Դասագիրքը նախատեսված է համալսարանների ֆիզիկամաթեմատիկական ֆակուլտետների ուսանողների և ասպիրանտների համար:

Տեսական հարուստ նյութից զատ, գիրքը աչքի է ընկնում նաև լուծված օրինակների բազմազանությամբ, ինչը հնարավորություն է տալիս դասագրքից օգտվելու նաև տեխնիկական և տնտեսագիտական բուհերի ուսանողներին:

ՀՏԴ 519.85(075.8)
ԳՄԴ 22.18ց73

ISBN 978-5-8084-2467-8

ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆ

Ն Ա Խ Ա Բ Ա Ն

Ընթերցողի ուշադրությանը ներկայացվող « Օպտիմիզացիայի մեթոդներ» դասագիրքը հայերենով համապարփակ և ծավալուն դասագրքի հրապարկման առաջին փորձն է: Այն ընդգրկում է համալսարանական ծրագրով նախատեսված օպտիմիզացիայի մեթոդներ առարկայի գրեթե բոլոր բաժինները:

Դասագիրքը ընդգրկում է Երևանի պետական համալսարանի ինֆորմատիկայի և կիրառական մաթեմատիկայի ֆակուլտետում հեղինակի կարդացած դասախոսությունները: Դասագրքի նպատակն է տալ օպտիմիզացիայի հիմնարար մեթոդների բավարար չափով համակարգված և ժամանակակից շարադրանք և ծանոթացնելու ընթերցողին այդ ալգորիթմների պրակտիկ իրականացումների հետ, որը այս գրքի կարևոր առանձնահատկություններից է:

Դասագիրքը աչքի է ընկնում շարադրանքի հստակությամբ և ապացույցների ճշտությամբ, լուծված օրինակների բազմազանությամբ: Նյութերը ներկայացվում են ամբողջական և բովանդակալից և տրվում են հասկացությունների երկրաչափական մեկնաբանություններ: Շարադրված մեթոդները պրակտիկ են և հեշտ իրագործելի ալգորիթմական լեզուների օգնությամբ:

Դասագիրքը բաղկացած է յոթ գլուխներից:

Առաջին գլխում համառոտակի շարադրվում են էքստրեմումի առաջին և երկրորդ կարգի անհրաժեշտ ու բավարար պայմանները ոչ պայմանական օպտիմիզացիայի խնդիրների համար վերջավոր չափանի տարածություններում:

Երկրորդ գլխում դիտարկվում են գրականության մեջ առավել հայտնի միաչափ օպտիմիզացիայի զրոյական և առաջին կարգի մեթոդները:

Երրորդ գլուխը նվիրված է պայմանական և ոչ պայմանական օպտիմիզացիայի մոտավոր մեթոդներին: Նկարագրվում են գրադիենտային մեթոդները և նրանցում քայլի ընտրության առավել հաճախ օգտագործվող կիսման, ապրիորի և ամենաարագ վայրէջքի եղանակները:

Չորրորդ գլխում ապացուցվում է Կուն- Տակկերի թեորեմը որպես մինիմումի անհրաժեշտ ու բավարար պայման ուռուցիկ ծրագրավորման խնդիրների համար: Այս գլխում տրվում են նաև որոշ հիմնարար թեորեմներ ուռուցիկ անալիզից, որոնք ունեն կիրառական լայն նշանակություն:

Հինգերորդ գլխում շարադրվում է մաթեմատիկական ծրագրավորման խնդիրների լուծման Լագրանժի անորոշ գործակիցների մեթոդը, երբ սահմանափակումները տրված են հավասարություններով և անհավասարություններով:

Վեցերորդ գլխում դիտարկվում են վարիացիոն հաշվի պարզագույն և իզոպերիմետրիկ խնդիրները և արտածվում է Էյլերի հավասարումը որպես էքստրեմումի անհրաժեշտ պայման:

Յոթերորդ գլխում շարադրվում է օպտիմալ կառավարման արագագործության խնդիրը գծային համակարգերում: Որպես օրինակ դիտարկվում է ֆիզիկական ճոճանակի օպտիմալ կառավարման խնդիրը:

Յուրաքանչյուր պարագրաֆի վերջում բերվում են խնդիրներ և վարժություններ, որոնց լուծումները ըն–թերցողին կօգնեն ավելի լավ յուրացնել շարադրված նյութը: Կան բարդ խնդիրներ, որոնք նշված են $*$ սիմվոլով, իսկ որոշ խնդիրների համար տրված են ցուցումներ:

Գրքում կան նաև առաջադրանքներ և խնդիրներ, որոնք կարող են լինել կուրսային և ավարտական աշխատանքներ ու իրականացվել համակարգչի օգնությամբ:

Վեցերորդ և յոթերորդ գլուխների նյութերը կարելի է օգտագործել մագիստրական կուրսերում կարդացվող վարիացիոն հաշվի և օպտիմալ կառավարման դասըն–թացներում:

Տեսության և խնդիրների դասակարգման լայն սպեկտրը հնարավորություն է տալիս դասագիրքը օգտագործել ոչ միայն ֆիզիկամաթեմատիկական ֆակուլտետներում, այլև տեխնիկական և տնտեսագիտական բուհերում:

Թեորեմի ապացույցը սկսվում է ► նշանով, իսկ ■ նշանը ազդարարում է ապացույցի ավարտը:

Շնորհակալություն ենք հայտնում ԵՊՀ թվային անա-լիզի և մաթեմատիկական մոդելավորման ամբիոնի աշ–խատակիցներին՝ ձեռնարկում ներկայացված նյութի բո-վանդակության և շարադրման եղանակների հետ կապված հարցերում օգտակար առաջարկությունների և դիտողու-թյունների համար:

Ռ.Ա. Խաչատրյան

ՀԻՄՆԱԿԱՆ ՆՇԱՆԱԿՈՒՄՆԵՐԻ ՑԱՆԿ

- $x \in X-$ x տարրը պատկանում է X բազմությանը
- $X \cap Y-$ բազմությունների հատում
- $X \cup Y-$ բազմությունների միավորում
- $X \backslash Y-$ բազմությունների տարբերություն
- $X + Y = \{z/z = x + y,\ x \in X,\ y \in Y\}-$ բազմությունների հանրահաշվական գումար
- $\overline{X}-$ X բազմության փակում
- $intX-$ X բազմության ներքին կետերի բազմություն
- R^n- n չափանի էվկլիդյան տարածություն
- $(x,y) = \sum_{i=1}^{n} x_i y_i-$ $x,\ y \in R^n$ վեկտորների սկալյար արտադրյալ
- $\|x\| = \sqrt{(x,x)}-$ $x \in R^n$ վեկտորի նորմ
- $B_r(a) = \{x \in R^n/\|x-a\| \leq r\}-$ a կենտրոնով r շառավղով գունդ
- $S_r(a) = \{x \in R^n/\|x-a\| = r\}-$ a կենտրոնով r շառավղով սֆերա
- $C^1[a,b]-$ $[a,b]$ հատվածի վրա որոշված անընդհատ դիֆերենցելի ֆունկցիաների տարածություն հետևյալ նորմով.

$$\|y(\cdot)\|_1 = \max\{\max_{x\in[a,b]} |y(x)|,\ \max_{x\in[a,b]} |y'(x)|\}$$

- $o(\alpha)-$ թվային ֆունկցիա, որը բավարարում է հետևյալ պայմանին.
$$\lim_{\alpha\to 0}\frac{o(\alpha)}{\alpha}=0$$
- A^T- A-ի տրանսպոնացված մատրիցը:

Գլուխ 1

Հիմնական գաղափարներ և թեորեմներ

Օպտիմիզացիայի խնդիրների ընդհանուր դրվածքը հետևյալն է: Տրված է որոշ M բազմություն և նրա վրա որոշված $f(x)$ ֆունկցիա: Պահանջվում է գտնել այդ ֆունկցիայի մինիմումի կամ մաքսիմումի կետերը այդ բազմության վրա: Այդ խնդիրը գրվում է հետևյալ կերպ.

$$f(x) \longrightarrow min(max),\ x \in M:$$

Այստեղ f-ը կոչվում է նպատակային ֆունկցիա, իսկ M-ը թույլատրելի բազմություն:

Դիցուք տրված է $f(x)$ ֆունկցիան R^n-ի վրա, և M-ը՝ ենթաբազմություն է R^n- ից:

Սահմանում 1.0.1: $x^* \in M$ *կետը կանվանենք*

1) f-ի գլոբալ մինիմումի (գլոբալ մաքսիմումի) կետ M բազմության վրա, եթե

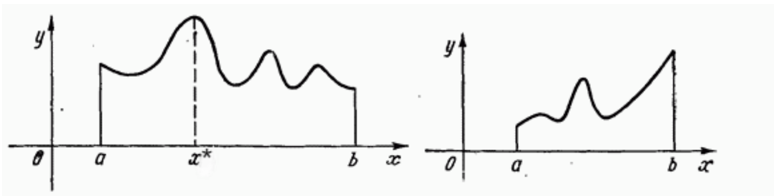
$$f(x) \geq f(x^*)\ (f(x) \leq f(x^*)) \quad \forall x \in M:$$

2) f-ի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ M բազմության վրա, եթե գոյություն ունի այդ կետի այնպիսի V շրջակայք, որ

$$f(x) \geq f(x^*) \; (f(x) \leq f(x^*)) \quad \forall x \in M \cap V:$$

Ֆունկցիայի մինիմումի և մաքսիմումի կետերը կոչվում են էքստրեմումի կետեր (տես գծ.1.1):

Օրինակ՝ առաջին գծապատկերում գլոբալ մաքսիմում ֆունկցիան հասնում է x^* կետում, իսկ երկրորդում ՝ b կետում: Նրանք ունեն նաև լոկալ և գլոբալ մինիմումներ, որոնք տարբեր են իրարից: Օպտիմիզացիայի խնդիրների



Գծ. 1.1: Ֆունկցիայի լոկալ և գլոբալ էքստրեմումներ

լուծման ժամանակ առաջին հերթին հարց է ծագում լուծման գոյության վերաբերյալ: Այդ տեսակետից հիշեցնենք հետևյալ կարևոր արդյունքը մաթեմատիկական անալիզից:

Թեորեմ 1.0.1 *(Վայերշտրաս): Դիցուք $M \subseteq R^n$ կոմպակտ բազմություն է, իսկ f-ը նրա վրա որոշված անընդհատ ֆունկցիա է: Այդ գոյություն ունեն f-ի գլոբալ մինիմումի և մաքսիմումի կետեր M-ի վրա:*

Այս գլխում բերվում են որոշ սահմանումներ և թեորեմներ ուռուցիկ անալիզից: Դիտարկվում է ողորկ ֆունկցիայի մի–նիմիզացիայի խնդիրը R^n էվկլիդյան տարածության վրա:

Շարադրվում են օպտիմալության առաջին ու երկրորդ կարգի անհրաժեշտ ու բավարար պայմանները:

Ենթադրվում է, որ ընթերցողը ծանոթ է մաթեմատիկական անալիզի և գծային հանրահաշվի հիմնարար գաղափարներին:

1.1 Նախնական սահմանումներ

Դիցուք $f(x) = f(x_1, x_2, \ldots, x_n)$ n փոփոխականի ֆունկցիա է՝ որոշված R^n էվկլիդյան տարածության վրա: Եթե f ֆունկցիան ըստ բոլոր փոփոխականների ունի մասնակի ածանցյալներ $x \in R^n$ կետում, ապա նրա գրադիենտը այդ կետում նշանակվում է հետևյալ կերպ.

$$f'(x) \equiv (f'_{x_1}(x), f'_{x_2}(x), ..., f'_{x_n}(x)):$$

Սահմանում 1.1.1: *Դիցուք $f(x)$-ը երկու անգամ դիֆերենցելի ֆունկցիա է $x \in R^n$ կետում:*
Հետևյալ սիմետրիկ մատրիցը կոչվում է հեսիան՝

$$H(x) = \begin{pmatrix} f''_{x_1x_1}(x) & f''_{x_1x_2}(x) & \ldots & f''_{x_1x_n}(x) \\ f''_{x_2x_1}(x) & f''_{x_2x_2}(x) & \ldots & f''_{x_2x_n}(x) \\ \ldots & \ldots & \ldots & \ldots \\ f''_{x_nx_1}(x) & f''_{x_nx_2}(x) & \ldots & f''_{x_nx_n}(x) \end{pmatrix}:$$

Դիցուք

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

կամայական մատրից է:

Սահմանում 1.1.2: **A** *մատրիցի* k*-րդ կարգի գլխավոր մինոր կոչվում է* $i_1 < i_2 < ... < i_k$ *համարներով տողերի և այդ նույն համարներով սյունների հատման տեղերում գտնվող տարրերից կազմված որոշիչը:*

Սահմանում 1.1.3: $\mathbf{A}(n \times n)$ *սիմետրիկ մատրիցը կոչ– վում է*

- *դրական որոշյալ* $(\mathbf{A} > 0)$*, եթե* $(\mathbf{A}x, x) > 0 \quad \forall x \neq \neq 0, \;\; x \in R^n$,
- *դրական կիսաորոշյալ* $(\mathbf{A} \geq 0)$*, եթե* $(\mathbf{A}x, x) \geq \geq 0 \;\; \forall x \in R^n$,
- *բացասական որոշյալ* $(\mathbf{A} < 0)$*, եթե* $(\mathbf{A}x, x) < < 0 \;\; \forall x \neq 0, \;\; x \in R^n$,
- *բացասական կիսաորոշյալ* $(\mathbf{A} \leq 0)$*, եթե* $(\mathbf{A}x, x) \leq \leq 0 \; \forall x \in R^n$*:*

Կարևոր կիրառական նշանակություն ունի հետևյալ պնդումը (տես, օրինակ՝ [27]):

Թեորեմ 1.1.1 *(Սիլվեստրի հայտանիշը)*:
Դիցուք $A(n \times n)$ *սիմետրիկ մատրից է:*

1) Որպեսզի **A** *մատրիցը լինի դրական որոշյալ, անհրաժեշտ է և բավարար, որ նրա գլխավոր անկյունագծային մինորները լինեն դրական՝*

$$\Delta_1 = a_{11} > 0, \; \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, ...,$$

$$\Delta_n = \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{vmatrix} > 0:$$

2) *Որպեսզի* $\mathbf{A}$ *մատրիցը լինի բացասական որոշյալ, անհրաժեշտ է և բավարար, որ* $\Delta_1 < 0, \ \Delta_2 > 0, ..., \ (-1)^n \Delta_n > 0$*:*

3) *Որպեսզի* $\mathbf{A}$ *մատրիցը լինի դրական կիսաորոշյալ, անհրաժեշտ է և բավարար, որ նրա գլխավոր մինորները լինեն ոչ բացասական:*

4) *Որպեսզի* $\mathbf{A}$ *մատրիցը լինի բացասական կիսաորոշյալ, անհրաժեշտ է և բավարար, որ զույգ կարգի գլխավոր մինորները լինեն ոչ բացասական, իսկ կենտ կարգի գլխավոր մինորները լինեն ոչ դրական:*

Սահմանում 1.1.4: $M \subseteq R^n$ *բազմությունը կոչվում է* ***ուռուցիկ***, *եթե ցանկացած* $x^1, \ x^2 \in M$ *կետերի և ցանկացած* $\alpha \in [0,1]$ *թվի համար տեղի ունի հետևյալը.*

$$\alpha x^1 + (1-\alpha)x^2 \in M:$$

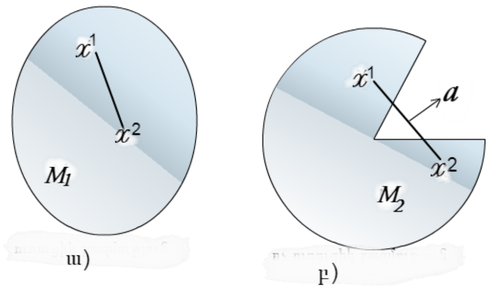
Սա նշանակում է, որ բազմությանը պատկանող երկու կետերը միացնող հատվածը ընկած է այդ նույն բազմության մեջ:

Գծ.1.2-ում պատկերված են ուռուցիկ և ոչ ուռուցիկ բազմությունների օրինակներ: M_2 բազմությունը ուռուցիկ չէ, որովհետև, օրինակ՝ a կետը M_2-ին չի պատկանում:

Թեորեմ 1.1.2: *Դիցուք* $M \subseteq R^n$ *փակ բազմություն է և կամայական* $x^1, x^2 \in M$ *կետերի համար գոյություն ունի այնպիսի* $\alpha \in (0,1)$ *թիվ, որ*

$$x_\alpha = \alpha x^1 + (1-\alpha)x^2 \in M:$$

Այդ դեպքում M*-ը ուռուցիկ է:*



Գծ. 1.2: ա) Ուռուցիկ բազմություն, բ) ոչ ուռուցիկ բազմություն

► Ենթադրենք հակառակը: Այդ դեպքում գոյություն կունենան այնպիսի $x^1, x^2 \in M$ կետեր և $\alpha_0 \in (0,1)$ թիվ, որ

$$y \equiv \alpha_0 x^1 + (1-\alpha_0)x^2 \notin M:$$

Նշանակենք

$$M_1 = M \cap [x^1, y],\ M_2 = M \cap [x^2, y]:$$

Այս բազմությունները կոմպակտ են, ընդ որում՝

$$y \notin M_1,\ y \notin M_2 :$$

Հետևաբար, կգտնվեն այնպիսի $y^1 \in M_1,\ y^2 \in M_2$ կետեր, որ

$$\min_{x \in M_1} \|x - y\| = \|y^1 - y\|,$$

$$\min_{x \in M_2} \|x - y\| = \|y^2 - y\|:$$

Ակնհայտ է, որ (y^1, y^2) միջակայքում չկան կետեր M բազմությունից, ինչը հակասում է թեորեմի պայմանին: ■

Թեորեմ 1.1.3: *Դիցուք $M \subseteq R^n$- ուռուցիկ բազմություն է և $x \in intM$, $y \in \overline{M}$: Այդ դեպքում*

$$(1-\alpha)x + \alpha y \in intM, \ \forall \alpha \in [0,1):$$

▶ Քանի որ $y \in \overline{M}$, ապա կամայական $\varepsilon > 0$ համար

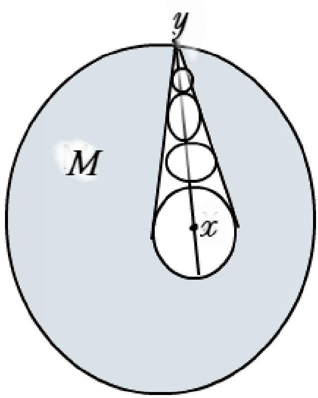
$$y \in M + B_\varepsilon(0):$$

Այդ դեպքում բավականաչափ փոքր $\varepsilon > 0$ համար կունենանք՝

$$(1-\alpha)x + \alpha y + B_\varepsilon(0) \subseteq (1-\alpha)x + \alpha(M + B_\varepsilon(0)) + B_\varepsilon(0) =$$

$$= (1-\alpha)[x + \varepsilon(1+\alpha)(1-\alpha)^{-1}B_1(0)] + \alpha M \subseteq$$
$$\subseteq (1-\alpha)M + \alpha M = M: \blacksquare$$

Երկրաչափորեն Թեորեմ 1.1.3-ի եզրակացությունը նշանակում է, որ $[x, y)$ կիսամիջակայքի բոլոր կետերը M բազմության ներքին կետեր են (տես գծ.1.3): Մենք այս հափկությունը համարում ենք ուռուցիկ բազմությունների հիմնական փոպոլոգիական հափկություն: Այս հափկությունից հեփևում է, որ եթե ուռուցիկ բազմությունը ունի ներքին կետեր, ապա նրա եզրային կետերը կարելի է մոտարկել ներքին կետերի միջոցով: Հեփագայում այս փասփը կարևոր նշանակություն կունենա ուռուցիկ բազմությունների բաժանման այն թորեմներում, որտեղ նրանք հափվում են եզրերով: Նշենք նաև Թեորեմ 1.1.3-ում M բազմության ուռուցիկության պայմանը էական է: Օրինակ, եթե $M = [0,1] \cap 2$, ապա $M = \overline{M}$, բայց $\overline{intM} = [0,1]$:



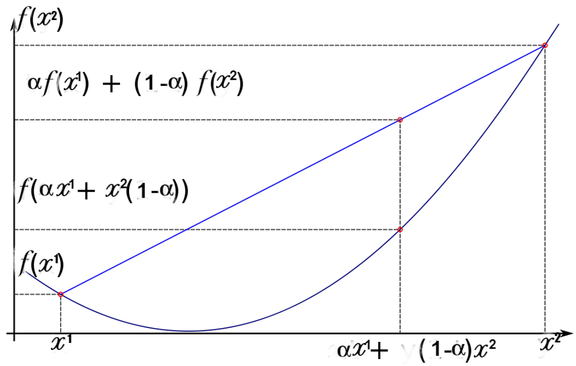
Գծ. 1.3: Ուռուցիկ բազմության հիմնական տոպոլոգիական հատկությունը. $[x, y) \subseteq intM$

Սահմանում 1.1.5: $f(x)$ *ֆունկցիան* $M \subseteq R^n$ *ուռուցիկ բազմության վրա կոչվում է* ***ուռուցիկ***, *եթե ցանկացած* $x^1,\ x^2 \in M$ *կետերի և ցանկացած* $\alpha \in [0, 1]$ *թվի համար տեղի ունի*

$$f(\alpha x^1 + (1 - \alpha)x^2) \leq \alpha f(x^1) + (1 - \alpha)f(x^2) \qquad (1.1.1)$$

անհավասարությունը:

Գծ.1.4-ում պատկերված է մեկ փոփոխականի ուռուցիկ ֆունկցիայի գրաֆիկ:



Գծ. 1.4: Ուռուցիկ ֆունկցիայի օրինակ

Օրինակ: Ցույց տանք, որ $f(x) = x^2$, $x \in R$ ֆունկցիան ուռուցիկ է:

Սպուգենք ուռուցիկ ֆունկցիայի անհավասարությունը՝

$$f(\alpha x_1 + (1-\alpha)x_2) = (\alpha x_1 + (1-\alpha)x_2)^2 =$$
$$= \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1 x_2 + (1-\alpha)^2 x_2^2 \leq$$
$$\leq \alpha^2 x_1^2 + \alpha(1-\alpha)(x_1^2 + x_2^2) + (1-\alpha)^2 x_2^2 =$$
$$= \alpha x_1^2 + (1-\alpha)x_2^2 = \alpha f(x_1) + (1-\alpha)f(x_2):$$

Թեորեմ 1.1.4: *Դիցուք f-ը ուռուցիկ ֆունկցիա է $M \subseteq R^n$ ուռուցիկ բազմության վրա և դիֆերենցելի է $x^* \in M$ կետում: Այդ դեպքում*

$$f(x) - f(x^*) \geq (f'(x^*), x - x^*) \quad \forall x \in M: \qquad (1.1.2)$$

► Ըստ ուռուցիկ ֆունկցիայի (1.1.1) սահմանման՝ կամայական $x \in M$ վեկտորի և ցանկացած $\alpha \in [0, 1]$ թվի համար ունենք

$$f(\alpha x + (1-\alpha)x^*) \leq \alpha f(x) + (1-\alpha)f(x^*)$$

անհավասարությունը: Եթե α-ն զրո չէ, ապա այս անհավա– սարությունը կարող ենք գրել հետևյալ տեսքով.

$$f(x) - f(x^*) \geq \frac{f(x^* + \alpha(x - x^*)) - f(x^*)}{\alpha}:$$

Այստեղ անցնենք սահմանի, երբ $\alpha \to 0$, և հաշվի առնելով, որ f ֆունկցիան դիֆերենցելի է x^* կետում, կստանանք

$$f(x) - f(x^*) \geq \lim_{\alpha \downarrow 0} \frac{f(x^* + \alpha(x - x^*)) - f(x^*)}{\alpha} =$$

$$= (f'(x^*), x - x^*): \blacksquare$$

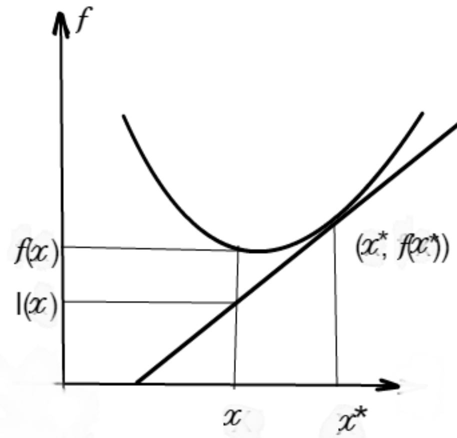
(1.1.2)-ը կոչվում է ուռուցիկ ֆունկցիայի **հիմնական անհավասարություն**: Հիշեցնենք, որ

$$l(x) = f(x^*) + (f'(x^*), x - x^*)$$

ֆունկցիայի գրաֆիկը կոչվում է շոշափող հիպերհարթություն՝ տարված f ֆունկցիայի գրաֆիկին $(x^*, f(x^*))$ կետում:

Հետևաբար (1.1.2) անհավասարությունը նշանակում է, որ եթե f ուռուցիկ ֆունկցիան դիֆերենցելի x^* կետում, ապա նրա գրաֆիկը գտնվում է $(x^*, f(x^*))$ կետով տարված շոշափող հիպերհարթությունից վերև (տես գծ .1.5): Կարելի է ցույց տալ, որ տեղի ունի նաև հակադարձ պնդումը. այսինքն՝ եթե ցանկացած $x^* \in R^n$ կետում տեղի ունի (1.1.2) անհավասարությունը, ապա f-ը ուռուցիկ է R^n-ի վրա (տես [24], Թեորեմ 3.6, էջ 89):

Այսպիսով, (1.1.2) անհավասարությունը կարելի է ընդունել որպես ուռուցիկ ֆունկցիայի սահմանում:



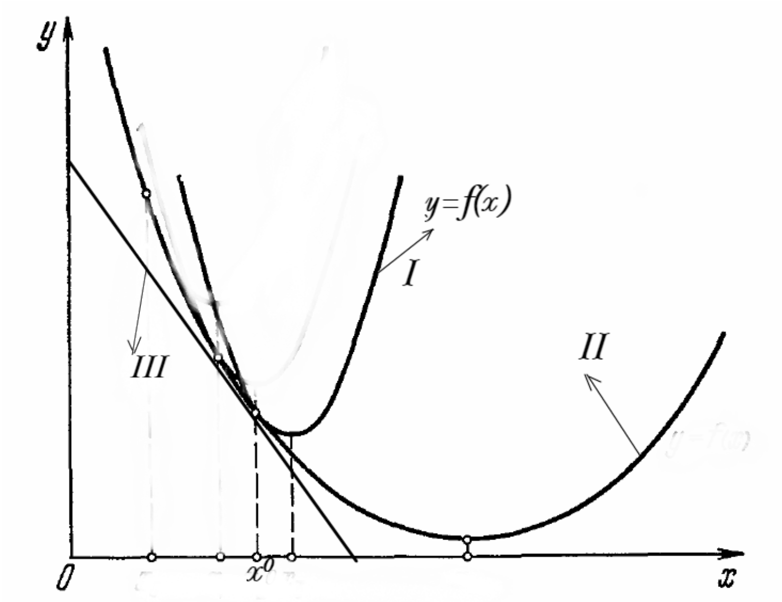
Գծ. 1.5: Դիֆերենցելի ուռուցիկ ֆունկցիայի գրաֆիկն ընկած է գրաֆիկին պարված շոշափող հիպերհարթությունից վերև:

Սահմանում 1.1.6: $f(x)$ *ֆունկցիան* $M \subseteq R^n$ *ուռուցիկ բազմության վրա կոչվում է* ***ուժեղ ուռուցիկ*** $\theta > 0$ *հաստատունով, եթե*

$$f(x^1) - f(x^2) \geq (f'(x^2), x^1 - x^2) + \theta \|x^1 - x^2\|^2 \quad \forall x^1, x^2 \in M:$$

Երկրաչափորեն սա նշանակում է, որ f ֆունկցիայի գրաֆիկի և նրա կամայական կետում պարված շոշափողի միջև կարելի է «նկարել» քառակուսային ֆունկցիայի գրաֆիկ (տես գծ.1.6): Այս սահմանումից հետևում է, որ եթե $f(x)$ ֆունկցիան ուժեղ ուռուցիկ է R^n վրա, ապա $f(x) \to +\infty$, երբ $\|x\| \to +\infty$:

Եթե ֆունկցիան երկու անգամ անընդհատ դիֆերենցելի է, ապա նրա ուռուցիկությունը ամբողջ փարածության վրա կարելի է ստուգել հեսիանի նշանի միջոցով: Այդ մասին է հետևյալ պնդումը (տես, օրինակ՝ [14], Թեորեմ 3.9, էջ 91):



Գծ. 1.6: Ուժեղ ուռուցիկ ֆունկցիայի երկրաչափական մեկնաբանությունը. I-ը $y = f(x)$ ֆունկցիայի գրաֆիկն է, II-ը ուժեղ ուռուցիկ ֆունկցիան է, իսկ III-ը $y = f(x)$ ֆունկցիայի գրաֆիկին $(x^0, f(x^0))$ կետում տարված շոշափողն է:

Թեորեմ 1.1.5: *Դիցուք f-ը երկու անգամ անընդհատ դիֆերենցելի է R^n-ի վրա: Այդ դեպքում*

ա) եթե $H(x) \geq 0 \ \forall x \in R^n$, ապա f-ը ուռուցիկ ֆունկցիա է R^n-ի վրա;

բ) եթե $(H(x)h, h) \geq \theta \|h\|^2 \ \forall x, h \in R^n$, ապա f-ը ուժեղ ուռուցիկ է θ հաստատունով R^n-ի վրա:

Այս թեորեմի կարևոր հետևանք է հետևյալ պնդումը:

Հետևանք 1.1.1: *Դիցուք* $A(n \times n)$*-ն դրական որոշյալ և սիմետրիկ մատրից է: Այդ դեպքում* $f(x) = 1/2(Ax, x) + (b, x)$ *ֆունկցիան ուժեղ ուռուցիկ է:*

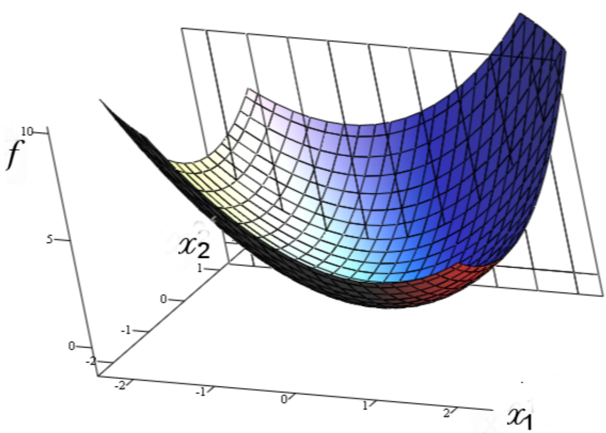
Իրոք, քանի որ $f''(x) = A$, ապա

$$\min_{x \in S_1(0)} (Ax, x) = \theta > 0:$$

Որպես օրինակ դիտարկենք $f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ ֆունկցիան: Այն կարող ենք ներկայացնել հետևյալ տեսքով. $f(x_1, x_2) \equiv f(x) = \frac{1}{2}(Ax, x)$, որտեղ

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}:$$

Քանի որ A մատրիցը սիմետրիկ է և դրական որոշյալ, ապա f ֆունկցիան ուժեղ ուռուցիկ է: Նրա գրաֆիկը ներկայացված է գծ.1.6-ում:



Գծ. 1.7: $f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ ուժեղ ուռուցիկ ֆունկցիայի գրաֆիկը

Օրինակ: Դիտարկենք $f(x) = e^x$ ֆունկցիան: Ունենք $f''(x) = e^x$: Այն ուժեղ ուռուցիկ է $[\alpha, \infty]$ բազմության վրա, որտեղ $\alpha-$ ն կամայական թիվ է: Իսկ $(-\infty, +\infty)$ բազմության վրա այն ուժեղ ուռուցիկ չէ, քանի որ $e^x \to 0$, երբ $x \to -\infty$:

Թեորեմ 1.1.6: *Ուռուցիկ բազմության վրա ուռուցիկ ֆունկցիայի լոկալ մինիմումի կետը հանդիսանում է նաև գլոբալ մինիմումի կետ:*

► Դիցուք x^*-ը f ուռուցիկ ֆունկցիայի լոկալ մինիմումի կետն է M ուռուցիկ բազմության վրա: Դա նշանակում է, որ գոյություն ունի x^* կետի այնպիսի $V(x^*)$ շրջակայք, որ

$$f(x) \geq f(x^*) \ \forall x \in M \cap V(x^*):$$

Բավականաչափ փոքր $\alpha > 0$ թվերի համար ունենք

$$\alpha x + (1-\alpha)x^* \in M \cap V(x^*):$$

Հետևաբար,

$$f(x^*) \leq f(\alpha x + (1-\alpha)x^*) \leq \alpha f(x) + (1-\alpha)f(x^*):$$

Այսպեղից $f(x^*) \leq f(x)$: Այսինքն՝ x^*-ը գլոբալ մինիմումի կետ է: ■

Թեորեմ 1.1.7: *Ուժեղ ուռուցիկ ֆունկցիան փակ ուռուցիկ բազմության վրա ունի միակ մինիմումի կետ այդ բազմության վրա:*

► Նախ ցույց տանք մինիմումի կետի գոյությունը: Դիցուք $x^0 \in M$ ֆիքսած կետ է: Դիտարկենք

$$P = \{x \in M / f(x) \leq f(x^0)\}$$

բազմությունը և ցույց փանք, որ այն սահմանափակ է: Ըստ ուժեղ ուռուցիկ Ֆունկցիայի սահմանման, եթե $x \in P$, ապա

$$x \in Q \equiv \{x \in R^n / \theta \|x - x^0\|^2 + (f'(x_0), x - x^0) \leq 0\}:$$

Հետևաբար, եթե ցույց փանք, որ Q-ն սահմանափակ բազմություն է, ապա սահմանափակ կլինի նաև P-ն: Ենթադրենք, որ Q-ն անսահմանափակ է: Դա նշանակում է, որ գոյություն ունի այնպիսի $\{x^i\} \in Q$ հաջորդականություն, որ $\|x^i\| \to +\infty$: Այստեղից կստանանք՝

$$\theta \|x^i - x^0\| + (f'(x^0), \frac{x^i - x^0}{\|x^i - x^0\|}) \leq 0: \qquad (1.1.3)$$

Նշանակենք՝

$$h^i = \frac{x^i - x^0}{\|x^i - x^0\|}:$$

Կարող ենք ենթադրել, որ $h^i \to h^0 \neq 0$:

(1.1.3) անհավասարության մեջ անցնելով սահմանի՝ կստանանք հակասություն, քանի որ

$$\theta \|x^i - x^0\| \to +\infty, \ (f'(x^0), \frac{x^i - x^0)}{\|x^i - x^0\|}) \to (f'(x^0), h^0):$$

Այսպիսով, քանի որ P-ն կոմպակտ բազմություն է, ապա f-ը հասնում է իր փոքրագույն արժեքին այդ բազմության վրա: Մյուս կողմից ակնհայտ է նաև, որ այդ արժեքը հենց ֆունկցիայի փոքրագույն արժեքն է ամբողջ տարածության վրա: Այժմ ցույց փանք մինիմումի կետի միակությունը: Դիցուք ենթադրենք, որ f-ը ունի երկու տարբեր մինիմումի կետեր՝ x^* և y^*: Ուժեղ ուռուցիկ ֆունկցիայի սահմանումից հետևում է, որ ցանկացած $x \neq x^*$ դեպքում $f(x) > f(x^*)$: Այս անհավասարության մեջ տեղադրելով $x = y^*$՝

կստանանք $f(y^*) > f(x^*)$, ինչը հակասություն է, քանի որ $f(x^*) = f(y^*)$:■

Թեորեմ 1.1.8: *Դիցուք $M \subseteq R^n$- ուռուցիկ կոմպակտ* [1] *է, իսկ $f(x)$-ը ուռուցիկ ֆունկցիա է՝ որոշված R^n-ի վրա: Եթե f-ը M-ի վրա հաստատունից տարբեր է, ապա նա այդ բազմության վրա հասնում է իր մեծագույն արժեքին միայն M-ի եզրային կետերում:*

▶ Ենթադրենք հակառակը: Դիցուք $B_r(x^*) \subseteq M$ և x^* կետում f ֆունկցիան ընդունում է մեծագույն արժեք: Ապացուցենք, որ $B_r(x^*)$ գնդի վրա f-ը հաստատուն է: Ենթադրենք գոյություն ունի այնպիսի $\overline{x} \in B_r(x^*)$, որ $f(\overline{x}) < f(x^*)$: Նշանակենք

$$\delta \equiv \|\overline{x} - x^*\|,\ \tilde{x} = x^* - \delta\frac{\overline{x} - x^*}{\|\overline{x} - x^*\|}\ :$$

Ունենք՝

$$f(x^*) = f(\frac{1}{2}(x^* + \delta\frac{\overline{x} - x^*}{\|\overline{x} - x^*\|}) + \frac{1}{2}(x^* - \delta\frac{\overline{x} - x^*}{\|\overline{x} - x^*\|})) \leq$$

$$\leq 1/2 f(\overline{x}) + 1/2 f(\tilde{x}) < 1/2 f(x^*) + 1/2 f(x^*) = f(x^*):$$

Ստացանք հակասություն:

Հետևաբար“

$$f(x) = f(x^*)\ \forall x \in B_r(x^*):$$

Այժմ ցույց տանք, որ $f(x) = f(x^*)$ $\forall x \in M$: Ենթադրենք հակառակը, որ գոյություն ունի $x \in M$ այնպիսին, որ $f(x) <$

[1] փակ և սահմանափակ

$f(x^*)$: Այդ դեպքում բավականաչափ փոքր դրական α թվերի համար
$$\alpha x + (1-\alpha)x^* \in B_r(x^*):$$
Այսպեղից, քանի որ նշված գնդի վրա f ֆունկցիան հաստատուն է, ապա
$$f(x^*) = f(\alpha x + (1-\alpha)x^*) \leq \alpha f(x) + (1-\alpha)f(x^*) <$$
$$< \alpha f(x^*) + (1-\alpha)f(x^*) = f(x^*),$$
ինչը հակասություն է: Այսպիսով ստացանք, որ f ֆունկցիան հաստատուն է M-ի վրա, որը հակասում է թեորեմի պայմանին:■

ԽՆԴԻՐՆԵՐ

---

1. Դիցուք M-ը ուռուցիկ բազմություն է:
   Ապացուցել, որ
$$(\alpha_1+\alpha_2)M = \alpha_1 M + \alpha_2 M \ \forall \alpha_1 \geq 0, \ \ \forall \alpha_2 \geq 0:$$

2. Արդյո՞ք հնարավոր է, որ երկու ոչ ուռուցիկ բազմությունների հանրահաշվական գումարը լինի ուռուցիկ (բերել համապատասխան օրինակներ):

3. Արդյո՞ք հնարավոր է, որ ուռուցիկ և ոչ ուռուցիկ բազմությունների հանրահաշվական գումարը լինի ուռուցիկ:

4. Դիցուք M-ը ուռուցիկ բազմություն է:
   Ապացուցել, որ

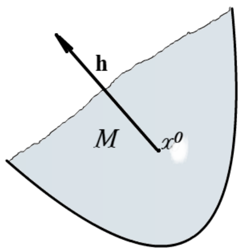
ա) $\overline{intM} = \overline{M}$,

բ) $\overline{M}$-ը ուռուցիկ է,

գ) $int\overline{M} = intM$:

Ցուցում: Օգտվել Թեորեմ 1.1.2-ից:

5*. Ապացուցել, որ երբ բազմությունը փակ է, անսահ– մանափակ և ուռուցիկ, ապա նրա կամայական կետով կարելի է տանել ճառագայթ, որն ամբողջովին ընկած կլինի այդ բազմության մեջ (տես գծ. 1.8):



Գծ. 1.8: Ճառագայթ անսահմանափակ ուռուցիկ բազմության մեջ

6. Ուսումնասիրել հետևյալ ֆունկցիայի ուռուցիկությունը.

$$f(x_1, x_2) = x_1^2 - 4x_1x_2 + 4x_2^2:$$

7. Ուսումնասիրել հետևյալ ֆունկցիայի ուռուցիկությունը.

$$f(x_1, x_2, x_3) = x_1x_3 - x_1^2 - x_2^2:$$

8. Ցույց տալ, որ

$$f(x_1, x_2) = \sqrt{1 + x_1^2 + x_2^2}$$

ֆունկցիան ուռուցիկ է R^2-ի վրա:

9. Նկարագրել բազմություն, որի վրա

$$f(x_1, x_2) = \frac{x_1^2}{x_2}$$

ֆունկցիան լինի ուռուցիկ:

10. a, b, c, պարամետրերի ինչպիսի՞ արժեքների դեպքում

$$f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2$$

ֆունկցիան կլինի ուռուցիկ R^2-ի վրա:

11. $f(x)$ ֆունկցիայի վերգրաֆիկ M ուռուցիկ բազմության վրա կոչվում է հետևյալ բազմությունը.

$$epi(f) \equiv \{(\alpha, x) \in R^{n+1} / x \in M, \;\; \alpha \geq f(x)\}:$$

Ապացուցել հետևյալ պնդումը: Որպեսզի f-ը լինի ուռուցիկ M ուռուցիկ բազմության վրա, անհրաժեշտ է և բավարար, որ նրա վերգրաֆիկը լինի ուռուցիկ բազմություն (տես գծ.1.9):

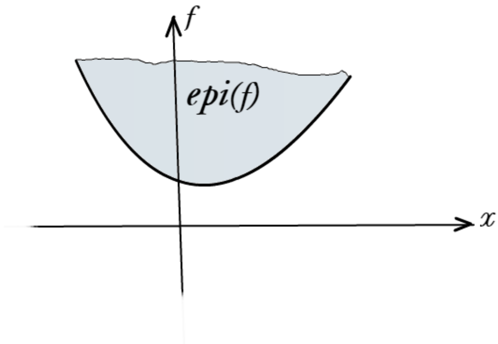
12. Դիցուք $f(x)$-ը ուռուցիկ ֆունկցիա է՝ որոշված M ուռուցիկ բազմության վրա և

$$x^i \in M, \; \alpha_i \geq 0, \; i \in [1:m], \; \sum_{i=1}^{m} \alpha_i = 1:$$

Ապացուցել, որ

$$f(\sum_{i=1}^{m} \alpha_i x^i) \leq \sum_{i=1}^{m} \alpha_i f(x^i):$$

Այս անհավասարությունը կոչվում է Յենսենի անհավասարություն:



Գծ. 1.9: Ստվերագծված մասը ֆունկցիայի վերգրաֆիկն է ($epi(f)$):

13. Դիցուք $f(x)$ ուռուցիկ ֆունկցիան սահմանափակ է վերևից R^n-ի վրա: Ապացուցել, որ f-ը հաստատուն է:

14*. Դիցուք $f(x)$ ուռուցիկ ֆունկցիան որոշված է $M \subseteq R^n$ բաց ուռուցիկ բազմության վրա: Ապացուցել, որ այն անընդհատ է այդ բազմության վրա:

15*. Դիցուք $f(x)$ ուռուցիկ ֆունկցիան դիֆերենցելի է $M \subseteq R^n$ բաց ուռուցիկ բազմության վրա: Ապացուցել, որ $f'(x)$ գրադիենտն անընդհատ է այդ բազմության վրա:

1.2 Էքստրեմումի առաջին և երկրորդ կարգի անհրաժեշտ ու բավարար պայմանները

Երբ օպտիմիզացիայի խնդրի թույլատրելի M բազմությունը համընկնում է ամբողջ տարածության հետ, ապա խնդիրը կոչվում է ոչ պայմանական օպտիմիզացիայի խնդիր.

$$f(x) \longrightarrow min(max),\ x \in R^n:$$

Այդ պայմանների կարևորությունը այն է, որ նրանք փալիս են էքստրեմումի կետերի բնութագիր և նրանց միջոցով հիմնավորվում են թվային այն մեթոդները, որոնք կառուցվում են օպտիմալ կետերը գտնելու համար: Ստորև ներկայացվում են առաջին և երկրորդ կարգի օպտիմալության պայմանները:

Թեորեմ 1.2.1 *(Էքստրեմումի առաջին կարգի անհրաժեշտ պայմանը)*: *Դիցուք x^*-ը $f(x)$ ֆունկցիայի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է R^n-ի վրա և f-ը դիֆերենցելի է այդ կետում:*

Այդ դեպքում f ֆունկցիայի գրադիենտը x^* կետում հավասար է զրոյի, այսինքն՝ $f'(x^*) = 0$, կամ, որ նույնն է՝

$$f'_{x_i}(x^*) = 0,\ i \in [1:n]:$$

► Քանի որ x^*-ը լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է, ապա օգտվելով ֆունկցիայի դիֆերենցելիության պայմանից, կամայական h վեկտորի և բավականաչափ փոքր α թվերի համար կունենանք

$$0 \underset{(\geq)}{\leq} f(x^* + \alpha h) - f(x^*) = (f'(x^*), \alpha h) + o(\alpha):$$

Բաժանելով այս անհավասարության երկու մասերը $\alpha > 0$ թվի վրա և ձգտեցնելով α-ն զրոյի՝ կստանանք

$$(f'(x^*), h) \geq 0\ ((f'(x^*), h) \leq 0)\ \forall h \in R^n:$$

Այստեղից անմիջականորեն հետևում է, որ $f'(x^*) = 0$: ■

Թեորեմ 1.2.2 *(Մինիմումի առաջին կարգի անհրաժեշտ ու բավարար պայմանը ուռուցիկ ֆունկցիայի համար)*:

Դիցուք f-ը ուռուցիկ ֆունկցիա է որոշված R^n-ի վրա *և դիֆերենցելի է x^* կետում: Որպեսզի x^*-ը լինի f-ի մինիմումի կետ R^n-ի վրա անհրաժեշտ է և բավարար, որ $f'(x^*) = 0$:*

► Անհրաժեշտությունը հետևում է **Թեորեմ 1.2.1-ից**: Ապացուցենք բավարարությունը: Օգտվելով ուռուցիկ ֆունկցայի հիմնական անհավասարությունից՝ ստանում ենք

$$f(x) - f(x^*) \geq (f'(x^*), x - x^*) = 0, \ \forall \, x \in R^n:$$

Այստեղից հետևում է, որ x^*-ը f-ի մինիմումի կետ է R^n-ի վրա: ■

Թեորեմ 1.2.3 *(Էքստրեմումի երկրորդ կարգի անհրաժեշտ պայմանը)*: *Դիցուք x^*-ը f ֆունկցիայի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է R^n-ի վրա և f-ը երկու անգամ դիֆերենցելի է այդ կետում:*

Այդ դեպքում $H(x^*)$-ը դրական կիսաորոշյալ է (բացասական կիսաորոշյալ է), այսինքն՝

$$H(x^*) \geq 0 \ \ (H(x^*) \leq 0):$$

► Քանի որ x^*-ը լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է, իսկ f-ը երկու անգամ դիֆերենցելի է այդ կետում, ապա կամայական $h \in R^n$ վեկտորի և բավականաչափ փոքր α թվերի համար ունենք

$$0 \underset{(\geq)}{\leq} f(x^* + \alpha h) - f(x^*) = 1/2(H(x^*)\alpha h, \alpha h) + o(\alpha^2):$$

Բաժանելով այս անհավասարության երկու մասերը α^2 թվի վրա և ձգտեցնելով α-ն զրոյի՝ կստանանք պահանջվող անհավասարությունը: ■

Թեորեմ 1.2.4 *(Էքստրեմումի երկրորդ կարգի բավարար պայմանները)*: *Դիցուք $f(x)$ ֆունկցիան երկու անգամ դիֆերենցելի է x^* կետում և տեղի ունեն հետևյալ պայմանները՝*

$$f'(x^*) = 0,\ H(x^*) > 0\ (H(x^*) < 0):$$

Այդ դեպքում x^*-ը f-ի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է R^n-ի վրա:

▶ Ենթադրենք հակառակը: Դա նշանակում է, որ գոյություն ունի այնպիսի $\{x^k\}$ հաջորդականություն, որ

$$x^k \to x^*,\ \ f(x^k) < f(x^*)\ \ (f(x^k) > f(x^*)) :$$

Նշանակելով $\alpha_k = \|x^k - x^*\|,\ \ h^k = (x^k - x^*)/\alpha_k$՝ կունենանք

$$x^k = x^* + \alpha_k h^k:$$

Քանի որ $\|h^k\| = 1$, ապա ընդհանրությունը չխախտելով կարող ենք ենթադրել, որ $h^k \to h^0 \neq 0$: Հաշվի առնելով թեորեմի ենթադրության $f'(x^*) = 0$ պայմանը՝ կունենանք

$$0 \underset{(\leq)}{\geq} f(x^k) - f(x^*) = 1/2(H(x^*)\alpha_k h^k, \alpha_k h^k) + o(\alpha_k^2):$$

Բաժանելով այս անհավասարության երկու մասերը α_k^2-ի վրա և անցնելով սահմանի՝ կստանանք

$$(H(x^*)h^0, h^0) \leq 0\ \ (H(x^*)h^0, h^0) \geq 0),$$

որը հակասում է թեորեմի ենթադրությանը: ■

Պարզագույն դեպքերում էքստրեմումի առաջին և երկրորդ կարգի անհրաժեշտ ու բավարար պայմանները էֆեկտիվ միջոցներ են ֆունկցիայի էքստրեմումի կետերը ճշգրիտ գտնելու համար:

Օրինակ: Գտնել

$$f(x) = x_1^3 + x_2^2 + x_3^2 + x_2x_3 - 3x_1 + 6x_2 + 2$$

ֆունկցիայի էքստրեմումի կետերը R^3-ի վրա:

Լուծում: Ըստ մինիմումի անհրաժեշտ պայմանի՝ ունենք

$$f'_{x_1} = 3x_1^2 - 3 = 0,\ f'_{x_2} = 2x_2 + x_3 + 6 = 0,\ f'_{x_3} = 2x_3 + x_2 = 0:$$

Լուծելով այս համակարգը՝ կստանանք երկու ստացիոնար կետ՝

$$x^1 = (1, -4, 2) \text{ և } x^2 = (-1, -4, 2):$$

Ունենք նաև, որ

$$f''_{x_1x_1} = 6x_1,\ f''_{x_1x_2} = 0,\ f''_{x_1x_3} = 0,$$

$$f''_{x_2x_2} = 2,\ f''_{x_2x_3} = 1,\ f''_{x_3x_3} = 2:$$

Այժմ յուրաքանչյուր ստացիոնար կետի համար կարելի է կազմել հեսիանը և ստուգել նրա նշանը: x^1 կետի համար հեսիանը ունի հետևյալ տեսքը.

$$H(x^1) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}:$$

Քանի որ

$$\Delta_1 = 6,\ \Delta_2 = \begin{vmatrix} 6 & 0 \\ 0 & 2 \end{vmatrix} = 12 > 0,\ \Delta_3 = 18 > 0,$$

ապա x^1-ը լոկալ մինիմումի կետ է: Ուսումնասիրենք x^2 կետը: Այդ կետում հեսիանը ունի հետևյալ տեսքը.

$$H(x^2) = \begin{pmatrix} -6 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}:$$

Քանի որ $\Delta_1 = -6 < 0$, $\Delta_2 = -12 < 0$, $\Delta_3 = -18 < 0$, ապա էքստրեմումի բավարար պայմանները տեղի չունեն: Ստուգենք երկրորդ կարգի անհրաժեշտ պայմանները: Առաջին կարգի գլխավոր մինորներն են՝ -6, 2, 2 թվերը: Երկրորդ կարգի գլխավոր մինորներն են՝ 3, -12, -12: Երրորդ կարգի գլխավոր մինորը հավասար է Δ_3-ի, որը բացասական է: Այսպիսով, x^2 կետում էքստրեմումի երկրորդ կարգի անհրաժեշտ պայմանները չեն կատարվում: Հետևաբար x^2 կետը էքստրեմումի կետ չէ:

ԽՆԴԻՐՆԵՐ

1. Գտնել $f(x)$ ֆունկցիայի էքստրեմումի կետերը.

$$f(x) = x_1^3 + x_2^2 + x_3^2 - 4x_1x_2 + 3x_1 + 2x_3 \to extr:$$

2. Ստուգել, արդյո՞ք $(1,1)$ կետը հետևյալ ֆունկցիայի էքստրեմումի կետ է, թե ոչ.

$$f(x) = (x_2 - x_1)^2 + (1 - x_1)^2 + 10(x_2 - 1)^2:$$

Գլուխ 2

Միաչափ օպտիմիզացիա

Սովորաբար n փոփոխականի $f(x_1, x_2, \ldots, x_n)$ ֆունկ–ցիայի մինիմիզացիայի իտերատիվ ալգորիթմներում անհրաժեշտ է գտնել մեկ փոփոխականի $\varphi(\alpha) = f(x + \alpha h)$ ֆունկցիայի մինիմումի կետը $[0, \infty)$ բազմության վրա, որտեղ h-ը f-ի նվազման ուղղությունն է, այսինքն՝ բավականաչափ փոքր α դրական թվերի համար

$$\varphi(\alpha) < \varphi(0):$$

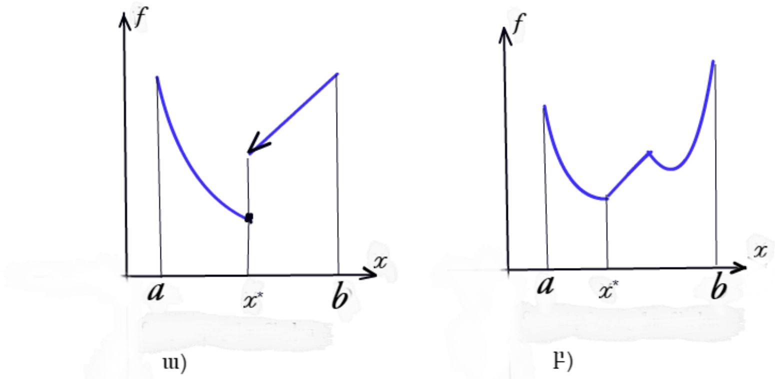
Սահմանում 2.0.1: *$f : R \to R$ ֆունկցիան կոչվում է ունիմոդալ*[1] *$[a, b]$ հատվածի վրա, եթե նա այդ հատվածի վրա ունի մինիմումի կետ՝ $x_* \in [a, b]$, և կամայական $\alpha, \beta \in [a, b]$, $\alpha < \beta$ թվերի համար տեղի ունեն հետևյալ անհավասարությունները.*

$$f(\alpha) > f(\beta), \ \beta \le x_*,$$

$$f(\alpha) < f(\beta), \ \alpha \ge x_*:$$

---

[1] Ֆունկցիա, որն ունի միակ էքստրեմում տրված միջակայքում

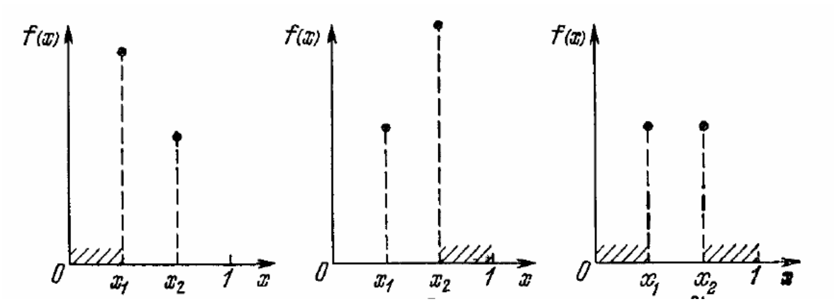


Գծ. 2.1: ա) ունիմոդալ ֆունկցիա, բ) ոչ ունիմոդալ ֆունկցիա

Այս սահմանումից հետևում է, որ ունիմոդալ ֆունկցիան ունի միակ մինիմումի կետ: Գծ.2.1-ում պատկերված են ունիմոդալ և ոչ ունիմոդալ ֆունկցիաների օրինակներ: Հեշտ է նկատել նաև, որ հատվածի վրա որոշված ուժեղ ուռուցիկ ֆունկցիան ունիմոդալ է:

Հայտնի է որ, f ֆունկցիայի մինիմումի կետը գտնելու համար պետք է լուծել $f'(x) = 0$ հավասարումը: Սակայն ֆունկցիան կարող է դիֆերենցելի չլինել և բացի դրանից՝ այդ հավասարման արմատների գտնելու խնդիրը կարող է լինել բավականաչափ բարդ: Այդ պատճառով կարևոր նշանակություն ունեն մինիմիզացիայի, այսպես կոչված, զրոյական մեթոդները, որոնցում անհրաժեշտ չէ իմանալ ֆունկցիայի ածանցյալը: Այդ մեթոդները սովորաբար կիրառվում են ունիմոդալ ֆունկցիաների մինիմիզացիայի ժամանակ: Ունիմոդալ ֆունկցիայի սահմանումից հետևում է, որ այն չի կարող ունենալ հասպապունության միջակայքեր: Ունիմոդալ ֆունկցիայի պարզագույն օրինակ է ուժեղ ուռուցիկ ֆունկցիան: Այն կարող Է ածանցյալ չունենալ և նույնիսկ անընդհատ չլինել: Ունիմոդալության հատկությունը թույլ է տալիս երկու փորձից հետո նշել մինիմումի կետը պարունակող ավելի փոքր ինտերվալ քան սկզբնականը: Այդ միջակայքերը կոչվում են անորոշության միջակայքեր: Իրոք, դիցուք $x_1 < x_2$: Այդ դեպքում հնարավոր են հետևյալ ելքերը.

1. $f(x_1) > f(x_2)$,
2. $f(x_1) < f(x_2)$,
3. $f(x_1) = f(x_2)$ (տես գծ.2.2):



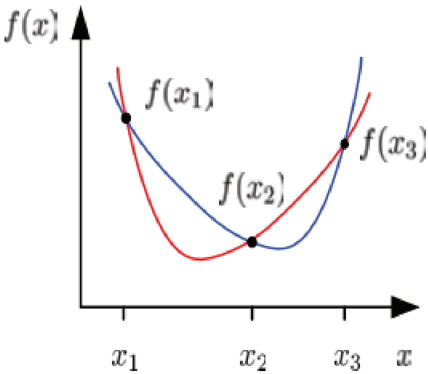
Գծ. 2.2: Անորոշության միջակայքը երկու փորձից հետո

Հեշտ է տեսնել, որ ֆունկցիայի մինիմումի կետը չի կարող գտնվել ստվերագծված ինտերվալներում: Հետևաբար, փորձերից հետո կախված ելքի արդյունքից կարելի է դեն գցել համապատասխան ինտերվալները: Մենք այստեղ դիտարկելու ենք միաչափ մինիմիզացիայի առավել հայտնի մեթոդները: Դրանք, այսպես կոչված, սիմետրիկ մեթոդներն են:

Այժմ ենթադրենք, f ֆունկցիան ունիմոդալ է ամբողջ թվային առանցքի վրա: Նկարագրենք մի ալգորիթմ, որը փալիս է մինիմումի կետը պարունակող $[a, b]$ միջակայք: Դա, այսպես ասած, «հաջող» եռյակի ընտրության մեթոդն է: x_1, x_2, x_3 կետերի եռյակը կոչվում է հաջող, եթե

$$x_1 < x_2 < x_3 \Longrightarrow f(x_1) > f(x_2) < f(x_3):$$

Հնարավոր է երկու տարբերակ, որ ֆունկցիան անցնի այդ կետերով: Ամեն դեպքում ֆունկցիայի ունիմոդալությունից հետևում է, որ նրա մինիմումը գտնվում է $[a = x_1, b = x_3]$ հատվածում (տես գծ.2.3): Հաջող եռյակ ընտրվում է հետևյալ



Գծ. 2.3: x_1, x_2, x_3 կետերի հաջող եռյակ

ալգորիթմով:

- Առաջին քայլում վերցվում է պատահական x_0 կետ և այդ կետի համար որոշվում է ֆունկցիայի նվազման ուղղությունը: Դրա համար վերցնում են կամայական h թիվ և հաշվում են ֆունկցիայի $f(x_0 + h)$ արժեքը: Եթե $f(x_0 + h) < f(x_0)$, ապա $x_1 = x_0 + h$ և անցում է կատարվում երկրորդ քայլին՝ $k = 1$ պայմանով : Եթե $f(x_0 + h) \geq f(x_0)$, ապա $h := -h$ և հաշվում են $f(x_0+h)$ արժեքը: Եթե, $f(x_0+h) < f(x_0)$, ապա ընդունում են $x_1 = x_0 + h$ և անցնում երկրորդ քայլին՝ ընդունելով $k = 1$: Եթե $f(x_0+h) \geq f(x_0)$, ապա ընդունում են $h := h/2$ և կրկնում են նախնական քայլի հաշվարկները: Առաջին քայլի ընթացքում սպացվում է այնպիսի h թիվ, որ $x_1 := x_0 + h$ և $f(x_1) < f(x_0)$:

- Երկրորդ քայլում h-ը կրկնապապտկվում է և ընդունում են $x_{k+1} = x_k + h$:

- Երրորդ քայլում հաշվվում են $f(x_{k+1})$ արժեքը: Եթե $f(x_{k+1}) < f(x_k)$, ապա ընդունում են $k := k + 1$ և անցնում են երկրորդ քայլին: Եթե $f(x_{k+1}) \geq f(x_k)$, ապա ալգորթմն ավարտվում է՝ ընդունելով $a = x_{k-1}$, $b = x_{k+1}$:

2.1 Զրոյական կարգի մեթոդներ

Ա) Փնտրման պասիվ մեթոդ

Այս մեթոդը ֆունկցիայի մինիմիզացիայի պարզագույն եղանակներից է: $[a, b]$ հատվածը տրոհում ենք k հավասար մասերի

$$x_i = a + i\frac{b-a}{k}, \quad i \in [0:k]$$

կետերով և գտնում ենք x_m կետն այնպես, որ

$$f(x_m) = \min_{i\in[0:k]} f(x_i):$$

x_m կետը համարվում է f ֆունկցիայի մինիմումի կետ: Ակն–հայտ է, որ պասիվ մեթոդով ֆունկցիայի մինիմումի կետը ϵ ճշտությամբ գտնելու համար անհրաժեշտ է $[a, b]$ հատվածը բաժանել $k \geq (b-a)/\epsilon$ հավասար մասերի:

Օրինակ: Դիցուք պետք է գտնել $f(x) = x + 2/x$ ֆունկցիայի մինիմումի կետը $[0.5, 3.5]$ հատվածի վրա $\varepsilon = 1/2$ ճշտությամբ: Ունենք

$$k \geq \frac{b-a}{\varepsilon} = 6 :$$

Հետևաբար

$$x_0 = 0.5,\ x_1 = 1,\ x_2 = 1.5,\ x_2 = 2,\ x_3 = 2.5,$$

$$x_4 = 2.5,\ x_5 = 3,\ x_6 = 3.5,$$

$$f(x_0) = 4.5,\ f(x_1) = 3,\ f(x_2) = 2.83,\ f(x_3) = 3,$$

$$f(x_4) = 3.3,\ f(x_5) = 3,\ f(x_6) = 4.07 :$$

Այսպեղից

$$x_m = x_2 = 1.5 :$$

Նշենք նաև, որ $x_* = \sqrt{2}$ -ը f ֆունկցիայի ճշգրիտ մինիմումի կետն է:

Բ) Դիխոտոմիայի (կիսման) մեթոդը δ **պարամետրով**: Այս մեթոդի յուրաքանչյուր իտերացիայում կառուցվում է մինիմումի կետը պարունակող հատված, որի երկարությունը փոքր է նախորդ իտերացիայում կառուցածից և համարյա երկու անգամ փոքրանում է անորոշության միջակայքը:

Կատարվում են հետևյալ քայլերը:

- Ընդունում ենք $a_1 = a,\ b_1 = b,\ \delta > 0,$

$$c_1 = \frac{a_1 + b_1}{2} - \frac{\delta}{2},\ d_1 = \frac{a_1 + b_1}{2} + \frac{\delta}{2}:$$

- Հաշվում ենք f ֆունկցիայի արժեքները c_1, d_1 կետերում և համեմատում իրար հետ: Եթե $f(c_1) \leq f(d_1)$, ապա $a_2 = a_1,\ b_2 = d_1$: Եթե $f(c_1) > f(d_1)$, ապա $a_2 = c_1,\ b_2 = b_1$ (տես գծ.2.4-ը):

- Այնուհետև, հաշվում ենք

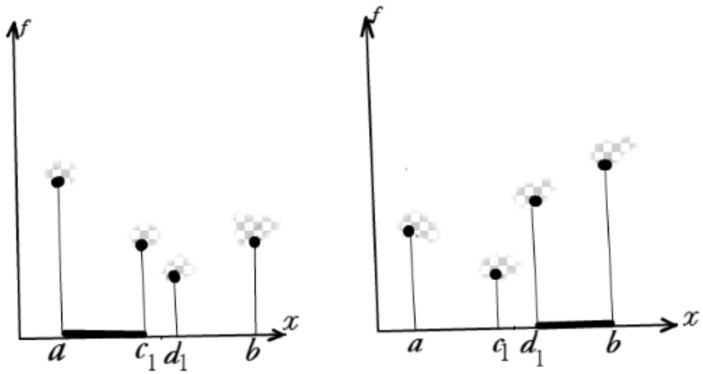
$$c_2 = \frac{a_2+b_2}{2} - \frac{\delta}{2},\ d_2 = \frac{a_2+b_2}{2} + \frac{\delta}{2}$$

արժեքները և համեմատում $f(c_2)$, $f(d_2)$ թվերը որոշելով նոր կետեր՝ a_3, b_3 և այսպես շարունակ մինչև որ տեղի ունենա

$$\varepsilon_i = \frac{b_i - a_i}{2} \leq \varepsilon$$

անհավասարությունը, որտեղ ε-ը նախապես տրված ճշտություն է: Իսկ f ֆունկցիայի մինիմումի կետ ε ճշտությամբ համարում ենք $x_m = (a_i + b_i)/2$ կետը:

Մեթոդի անվանումը կապված է այն բանի հետ, որ փոքր $\delta > 0$ թվերի համար մինիմումի կետը տեղայնացնող (լոկալիզացնող) միջակայքը փոքրանում է համարյա երկու անգամ յուրաքանչյուր քայլից հետո: Սակայն նկատենք, որ բավականաչափ փոքր δ-ի ընտրության դեպքում պետք է f ֆունկցիայի $f(c_i)$, $f(d_i)$ արժեքները հաշվել մեծ ճշգրտությամբ: Որովհետև հնարավոր է այնպիսի դեպք, որ ֆունկցիայի մինիմումի կետը գտնվի (d_i, b_i) միջակայքում, բայց ֆունկցիայի արժեքների ոչ ճշգրիտ հաշվարկի դեպքում ստացվի, որ $f(c_i) = f(d_i)$ (բայց իրականում $f(c_i) > f(d_i)$): Եվ այդ ժամանակ նոր տեղայնացնող միջակայքը կլինի $[a_i, d_i]$ հատվածը, որը արդեն չի պարունակի մինիմումի կետը:



Գծ. 2.4: Ֆունկցիայի մինիմումի կետը պեղայնացնող հատվածի (անորոշության միջակայքի) կառուցումը դիխոտոմիայի մեթոդով

Օրինակ: Գտնենք $f(x) = x+2/x$ ֆունկցիայի մինիմումի կետը $[0.5,\ 3.5]$ հատվածի վրա $\varepsilon = 0.5$ ճշտությամբ և դիցուք $\delta = 0.1$:

Կիրառելով դիխոտոմիայի մեթոդի քայլերը՝ կունենանք հետևյալ արդյունքները:

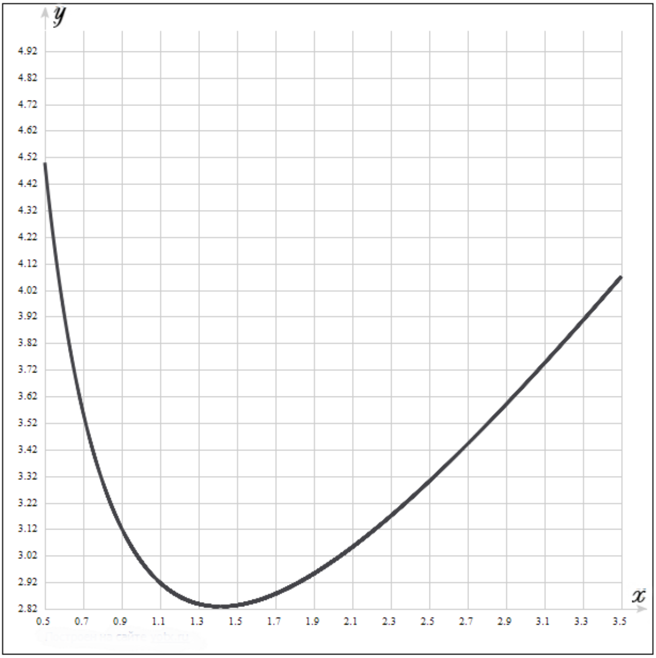
- Առաջին քայլում $a_1 = 0.5,\ b_1 = 3.5,$

$$c_1 = \frac{0.5+3.5}{2} - \frac{0.1}{2} = 1.95,\ d_1 = \frac{0.5+3.5}{2} + \frac{0.1}{2} = 2.05,$$

$$f(c_1) = 2.976 < f(d_1) = 3.026:$$

- Երկրորդ քայլում $a_2 = a_1 = 0.5,\ b_2 = d_1 = 2.05,$

$$\varepsilon_2 = \frac{2.05-0.5}{2} = 0.775 > 0.5 :$$



Գծ. 2.5: $y = x + 2/x$ Ֆունկցիայի գրաֆիկը

Ուրեմն կանգառի պայմանը տեղի չունի. հետևաբար իտերացիան շարունակում ենք.

$$c_2 = \frac{0.5 + 2.05}{2} - \frac{0.1}{2} = 1.225,$$

$$d_2 = \frac{0.5 + 2.05}{2} + \frac{0.1}{2} = 1.325,$$

$$f(c_2) = 2.858 > f(d_2) = 2.834:$$

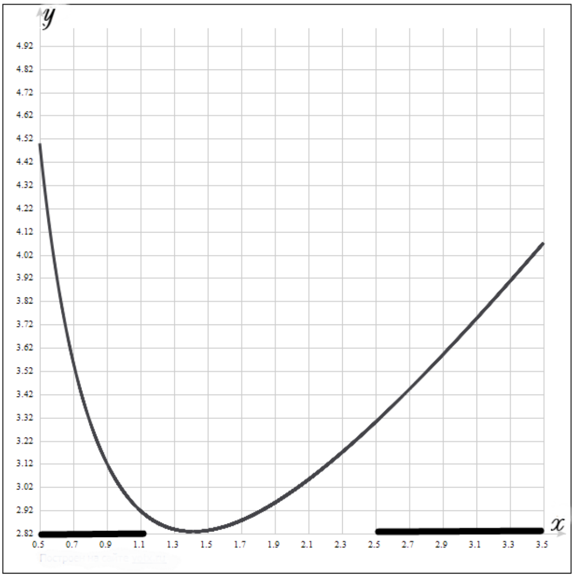
- Երրորդ քայլում $a_3 = c_2 = 1.225, b_3 = d_2 = 2.05,$

$$\varepsilon_3 = \frac{2.05 - 1.225}{2} = 0.4125 < 0.5:$$

Ինչպես տեսնում ենք կանգառի պայմանը կատարվում է, հետևաբար իտերացիան ավարտում ենք ընդունելով՝

$$x_m = \frac{2.05 + 1.225}{2} = 1.638։$$

Գծ.2.6-ում պատկերված են դիխոտոմիայի մեթոդի յուրաքանչյուր իտերացիայում դեն նետված հատվածները։



Գծ. 2.6: Անորոշության միջակայքերը $y = x + 2/x$ ֆունկցիայի համար

Նշենք, որ այս մեթոդով մինիմումի կետը $\varepsilon = 0.5$ ճշտությամբ ստանալու համար անհրաժեշտություն եղավ հաշվել f ֆունկցիայի արժեքները վեց կետերում։ Իսկ պասիվ մեթոդով նույն ճշտությունը ապահովելու համար պետք էր հաշվել f-ի արժեքները յոթ կետերում:

Գ) Ոսկե հատման մեթոդը: Դիխոտոմիայի մեթոդի յուրաքանչյուր իտերացիայում անհրաժեշտ էր հաշվել ֆունկցիայի արժեքները երկու c_i, d_i նոր կետերում: Ոսկե հատման մեթոդի յուրաքանչյուր իտերացիայում հաշվվում են ֆունկցիայի արժեքը միայն մեկ նոր կետում և դրանով իսկ ավելի քիչ անգամ են դիմում ֆունկցիայի արժեքները հաշվող ծրագրին:

Կասենք, որ $D \in [A, B]$ կետով կատարվում է $[A, B]$ հատվածի ոսկե հատում, եթե

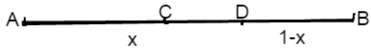
$$\frac{AB}{AD} = \frac{AD}{DB} \text{ (տես գծ. 2.7):}$$

Դիցուք $AB = 1$ և $AD = x$: Այդ դեպքում

$$\frac{1}{x} = \frac{x}{1-x}:$$

Այստեղից

$$x = \frac{\sqrt{5}-1}{2}:$$



Գծ. 2.7: Հատվածի ոսկե հատումը

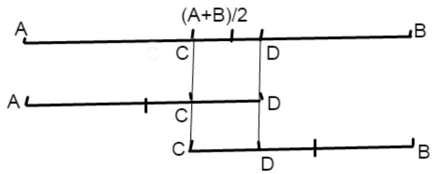
Կենտրոնի նկատմամբ D-ին համաչափ կետը նշանակենք C-ով: Ակնհայտ է, որ

$$AC = \frac{3-\sqrt{5}}{2}:$$

C կետը կոչվում է ոսկե հատման առաջին կետ, իսկ D-ն՝ երկրորդ: Այդ կետերը օժտված են հետևյալ հրաշալի հատկությամբ.

- C-ն ոչ միայն $[A, B]$ հատվածի ոսկե հատման առաջին կետն է, այլև՝ $[A, D]$ հատվածի ոսկե հատման երկրորդ կետը:

- D-ն ոչ միայն $[A, B]$ հատվածի ոսկե հատման երկրորդ կետն է, այլև՝ $[C, B]$ հատվածի ոսկե հատման առաջին կետը (տես գծ.2.8):



Գծ. 2.8: Հատվածի ոսկե հատման հրաշալի հատկությունը

Ոսկե հատման մեթոդի յուրաքանչյուր իտերացիա հաշվում է ֆունկցիայի արժեքը միայն մեկ նոր կետում:
Նկարագրենք ոսկե հատման մեթոդի ալգորիթմը:

- Ընդունում ենք $a_1 = a,\ b_1 = b$ և հաշվում՝

$$c_1 = \frac{3-\sqrt{5}}{2}(b_1 - a_1) + a_1,\ d_1 = \frac{\sqrt{5}-1}{2}(b_1 - a_1) + a_1:$$

- Եթե $f(c_1) \leq f(d_1)$, ապա ընդունում ենք, որ

$$a_2 = a_1, b_2 = d_1,\ d_2 = c_1,$$

$$c_2 = \frac{3-\sqrt{5}}{2}(b_2 - a_2) + a_2:$$

Եթե $f(c_1) > f(d_1)$, ապա՝ $a_2 = c_1,\ b_2 = b_1,\ c_2 = d_1,$

$$d_2 = \frac{\sqrt{5}-1}{2}(b_2 - a_2) + a_2:$$

- Համեմատելով $f(c_2)$ և $f(d_2)$ արժեքները՝ որոշում ենք a_3, b_3-ի նոր արժեքները և այսպես շարունակ մինչև որ տեղի ունենա

$$\varepsilon_i = \frac{b_i - a_i}{2} \leq \varepsilon$$

անհավասարությունը, որտեղ ε-ը անհրաժեշտ ճշտությունն է:

Յուրաքանչյուր քայլում մինիմումի կետը տեղայնացնող միջակայքը փոքրանում է $\frac{\sqrt{5}-1}{2}$ գործակցով:
Հետևաբար

$$\varepsilon_i = \frac{1}{2}(\frac{\sqrt{5}-1}{2})^{i-1}(b-a):$$

Օրինակ: Դիտարկենք նախորդ օրինակի ֆունկցիան.

$$f(x) = x + \frac{2}{x},\ a = 0.5,\ b = 3.5,\ \varepsilon = 0.5:$$

Կիրառելով ոսկե հատման մեթոդի ալգորիթմը՝ կունենանք հետևյալ արդյունքները:

- Առաջին քայլում $a_1 = 0.5, b_1 = 3.5,$

$$c_1 = \frac{3-\sqrt{5}}{2}(3.5-0.5)+0.5 = 1.646,$$
$$d_1 = \frac{\sqrt{5}-1}{2}(3.5-0.5)+0.5 = 2.354,$$
$$f(c_1) = 2.861 < f(d_1) = 3.204:$$

- Երկրորդ քայլում $a_2 = a_1 = 0.5, \quad b_2 = d_1 = 2.354,\ d_2 = c_1 = 1.646,$

$$c_2 = \frac{3-\sqrt{5}}{2}(b_2-a_2)+a_2 = 1.208,$$

$$\varepsilon_3 = \frac{2.354-0.5}{2} = 0.927 > 0.5:$$

Հետևաբար, ալգորիթմը շարունակում ենք, քանի որ կանգառի պայմանը չի կատարվում. $f(c_2) = 2.864 > f(d_2) = 2.861:$

- Երրորդ քայլում $a_3 = c_2 = 1.208,\ b_3 = b_2 = 2.354, c_3 = d_2 = 1.646,$

$$d_3 = \frac{\sqrt{5}-1}{2}(b_3-a_3)+a_3 = 0.573 = 1.916,$$

$$\varepsilon_3 = \frac{2.354-1.208}{2} = 0.573 > 0.5:$$

Ուրեմն կանգառի պայմանը այս քայլում նույնպես տեղի չունի, ուստի պրոցեսը շարունակում ենք.

$$f(c_3) = f(d_2) = 2.861 < f(d_3) = 2.96:$$

- Չորրորդ քայլում $a_4 = a_3 = 1.208$, $b_4 = d_3 = 1.916$, $d_4 = c_3 = 1.646$,

$$c_4 = \frac{3-\sqrt{5}}{2}(b_4 - a_4) + a_4 = 1.478,$$

$$\varepsilon_4 = \frac{1.916 - 1.208}{2} = 0.354 < 0.5:$$

Ինչպես տեսնում ենք այս քայլում արդեն կանգառի պայմանը տեղի ունի, ուստի ընդունում ենք

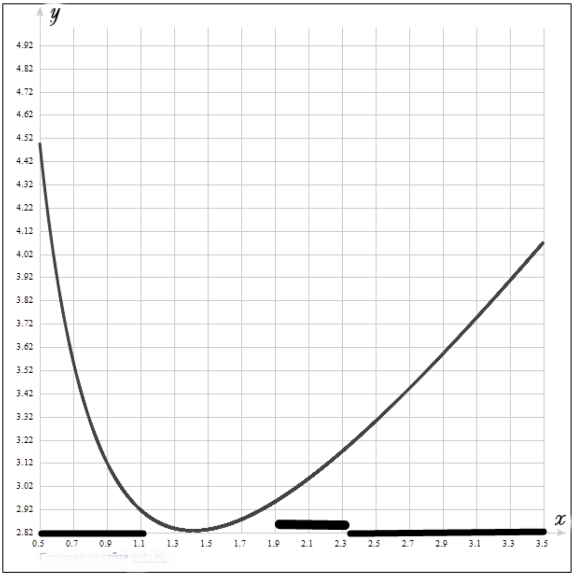
$$x_m = \frac{1.916 + 1.208}{2} = 1.562$$

և ալգորիթմը ավարտվում է: Գծ.2.9-ում պատկերված են ոսկե հատման մեթոդով յուրաքանչյուր իտերացիայում դեն նետված հատվածները: Նշենք, որ այստեղ տրված ճշտությունը ապահովելու համար անհրաժեշտ եղավ դիմել ֆունկցիայի արժեքները հաշվող ծրագրին հինգ անգամ, այսինքն՝ ավելի քիչ քան դիխոտոմիայի մեթոդում:

Նշենք, որ ոսկե հատման մեթոդի յուրաքանչյուր իտերացիայից հետո անորոշության միջակայքը կրճատվում է $\beta = (\sqrt{5}-1)/2$ անգամ: Հետևաբար, ϵ-ճշտությամբ մինիմումի կետը ստանալու համար անհրաժեշտ իտերացիաների N քանակը կարելի է որոշել

$$\epsilon < (b-a)\beta^N$$

անհավասարությունից: Արդյունավետության տեսակետից ոսկե հատման մեթոդը միջանկյալ դիրք է գրավում դիխոտոմիայի և Ֆիբոնաչիի մեթոդների միջև: Ֆիբոնաչիի մեթոդը կշարադրվի ստորև:



Գծ. 2.9: Հատվածի ոսկե հատման մեթոդով դեն նետված միջակայքերը $y = x + 2/x$ ֆունկցիայի համար

Դ) Ֆիբոնաչիի մեթոդը:

Ֆիբոնաչիի հաջորդականությունը կառուցվում է հետևյալ ռեկուրենտ առնչությամբ՝

$$F_1 = 1,\ F_2 = 1,\ F_{n+2} = F_n + F_{n+1},\ n \geq 1 :$$

Դուրս բերենք Ֆիբոնաչիի հաջորդականության F_n ընդհանուր անդամի բանաձևը: x-ը ընտրենք այնպես, որ x^n հաջորդականությունը բավարարի Ֆիբոնաչիի հաջորդականության ռեկուրենտ առնչությանը:
Ունենք

$$x^{n+2} = x^n + x^{n+1}:$$

Այստեղից հետևում է, որ x-ը պետք է բավարարի հետևյալ քառակուսի հավասարմանը.

$$x^2 - x - 1 = 0:$$

Այս հավասարման արմատներն են՝

$$\tau = \frac{1+\sqrt{5}}{2} \cong 1.618, \quad -\frac{1}{\tau} = \frac{1-\sqrt{5}}{2}:$$

Հետևաբար, τ^n, $(\frac{-1}{\tau})^n$ հաջորդականությունները բավարարում են Ֆիբոնաչիի հաջորդականության ռեկուրենտ առնչությանը: Ակնհայտ է, որ այդ առնչությանը բավարարում է նաև

$$c_1\tau^n + c_2(-\frac{1}{\tau})^n$$

հաջորդականությունը, որտեղ $c_1,\ c_2$-ը կամայական թվեր են: $c_1,\ c_2$ հաստատունները ընտրենք այնպես, որ բավարարվեն $F_1 = 1,\ F_2 = 1$ սկզբնական պայմանները:
Այստեղից

$$c_1 + c_2 = 1,\ c_1\tau + c_1(-\frac{1}{\tau}) = 1:$$

Հետևաբար Ֆիբոնաչիի հաջորդականության ընդհանուր անդամը կունենա հետևյալ տեսքը.

$$F_n = \frac{\tau^{n+1} - \left(-\tau\right)^{-(n+1)}}{\sqrt{5}} = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}}{\sqrt{5}}:$$

Այստեղից կստանանք

$$F_n \sim \frac{\tau^{n+1}}{\sqrt{5}}:$$

Հետևաբար

$$\lim_{n\to\infty} \frac{F_n}{F_{n+1}} = \frac{\sqrt{5}-1}{2}:$$

Նշանակենք

$$c_1 = a_1 + (b_1 - a_1)\frac{F_n}{F_{n+2}},$$

$$d_1 = a_1 + (b_1 - a_1)\frac{F_{n+1}}{F_{n+2}},\ n > 1:$$

Այս երկու կետերը կենտրոնի նկատմամբ համաչափ են դասավորված: Իրոք, ունենք

$$c_1 - a_1 = (b_1 - a_1)\frac{F_n}{F_{n+1}} = (b_1 - a_1)(1 - \frac{F_{n+1}}{F_{n+2}}) = b_1 - d_1,$$

ինչը նշանակում է նշված կետերի համաչափ դասավորվածություն կենտրոնի նկատմամբ: Այնուհետև համեմատում ենք $f(c_1)$ և $f(d_1)$ արժեքները՝ դեն նետելով $[a_1, c_1)$ $[d_1, b_1)$ կիսամիջակայքերից որևէ մեկը: Եթե $f(c_1) < f(d_1)$, ապա դեն ենք նետում $(d_1, b_1]$-ը: Այս դեպքում c_1 կետը մինիմումի կետը պեղայնացնող $[a_2, b_2] \equiv [a_1, d_1]$ հատվածը բաժանում է

$$\frac{c_1 - a_1}{d_1 - a_1} = \frac{\frac{F_n}{F_{n+2}}}{\frac{F_{n+1}}{F_{n+2}}} = \frac{F_n}{F_{n+1}}$$

հարաբերությամբ: Ակնհայտ է նաև, որ

$$d_2 \equiv a_2 + (b_2 - a_2)\frac{F_n}{F_{n+1}} = c_1 :$$

Եթե $f(c_1) > f(d_1)$, ապա դեն ենք նետում $[a_1, c_1)$ միջակայքը: d_1 կետով մինիմումի կետը լոկալիզացնող $[a_2, b_2] \equiv [c_1, b_1]$ հատվածը բաժանվում է

$$\frac{d_1 - c_1}{b_1 - c_1} = \frac{d_1 - c_1}{d_1 - a_1} = \frac{\frac{F_{n+1}}{F_{n+2}} - \frac{F_n}{F_{n+2}}}{\frac{F_{n+1}}{F_{n+2}}} = \frac{F_{n-1}}{F_{n+1}} :$$

հարաբերությամբ: Հետևաբար այս դեպքում

$$c_2 \equiv a_2 + (b_2 - a_2)\frac{F_{n-1}}{F_{n+1}} = d_1$$

Այսպիսով, ընդունելով $a_1 = a$, $b_1 = b$, կառուցում ենք $\{c_i\}, \{d_i\}$ հաջորդականությունները հետևյալ բանաձևերով.

$$c_i \equiv a_i + (b_i - a_i)\frac{F_{n+1-i}}{F_{n+3-i}},$$

$$d_i \equiv a_i + (b_i - a_i)\frac{F_{n+2-i}}{F_{n+3-i}} :$$

Նշենք, որ յուրաքանչյուր իտերացիայում հաշվում ենք ֆունկցիայի արժեքը միայն մեկ նոր կետում (c_i կամ d_i) և բացի դրանից

$$b_2 - a_2 = (b_1 - a_1)\frac{F_{n+1}}{F_{n+2}},$$

$$b_3 - a_3 = (b_2 - a_2)\frac{F_n}{F_{n+1}} = (b_1 - a_1)\frac{F_n}{F_{n+1}}\frac{F_{n+1}}{F_{n+2}} = (b_1 - a_1)\frac{F_n}{F_{n+2}} :$$

Հետևաբար

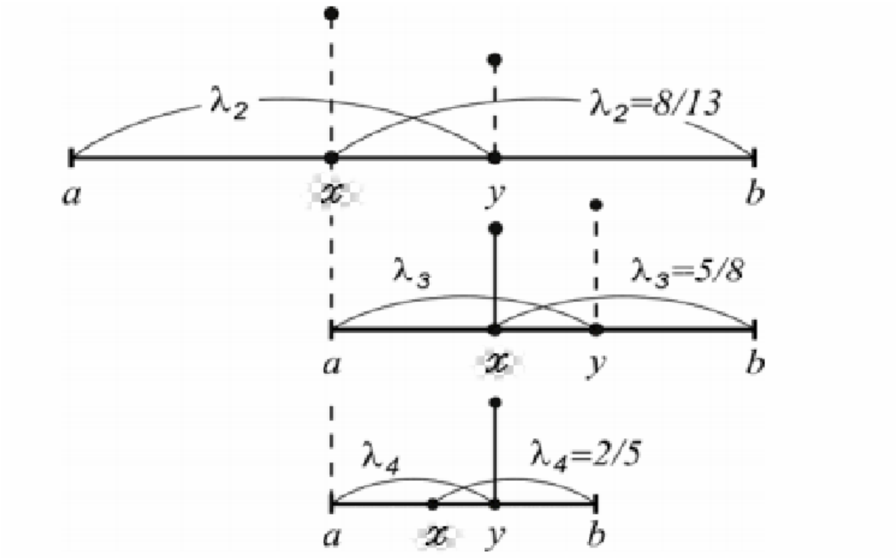
$$b_n - a_n = (b_1 - a_1)\frac{F_3}{F_{n+2}}:$$

Այստեղից հետևում է, որ մինիմումի կետը ε ճշտությամբ վերականգնելու համար անհրաժեշտ n քայլերի քանակը պետք է բավարարի հետևյալ անհավասարությանը.

$$(b - a)\frac{F_3}{F_{n+2}} \le 2\varepsilon,$$

այսինքն՝

$$F_{n+2} \ge \frac{b - a}{\varepsilon}:$$



Գծ. 2.10: Ֆիբոնաչիի մեթոդը $N = 6$ դեպքում: Նշված են նաև յուրաքանչյուր քայլում դեն նետված հատվածների չափերը:

Օրինակ: Դիցուք

$$f(x) = x + \frac{2}{x},\ a = 0.5,\ b = 3.5,\ \varepsilon = 0.5:$$

Ֆիբոնաչիի մեթոդով գտնենք f ֆունկցիայի մինիմումի կետը $[a, b]$ հատվածի վրա ε ճշտությամբ:

Ունենք $F_{n+2} \geq \frac{b-a}{\varepsilon} = 6$: Այստեղից $n \geq 4$, այսինքն՝ Ֆիբոնաչիի մեթոդով պետք է կատարել չորս քայլ մինիմումի կետը ε ճշտությամբ սպանալու համար: Կիրառելով Ֆիբոնաչիի մեթոդի ալգորիթմը՝ կունենանք հետևյալ արդյունքները:

- Առաջին քայլում $n = 4,\ a_1 = 0.5,\ b_1 = 3.5,$

$$c_1 = 0.5 + (3.5 - 0.5)\frac{F_4}{F_6} = 0.5 + 3 \cdot \frac{3}{8} = 1.625,$$

$$d_1 = a_1 + (3.5 - 0.5)\frac{F_5}{F_6} = 0.5 + 3 \cdot \frac{5}{8} = 2.375,$$

$$f(c_1) = 2.856 < f(d_1) = 3.217:$$

- Երկրորդ քայլում $a_2 = a_1 = 0.5,\ b_2 = d_1, = 2.375,\ d_2 = c_1,$

$$c_2 = 0.5 + (2.375 - 0.5)\frac{F_3}{F_5} = 0.5 + 1.875 \cdot \frac{3}{5} = 1.25,$$

$$f(c_2) = 2.85,\ f(d_2) = f(c_1) = 2.856,$$

$$f(c_2) < f(d_2):$$

- Երրորդ քայլում $a_3 = a_2 = 0.5,\ b_3 = d_2 = 1.625,\ d_3 = c_2,$

$$c_3 = 0.5 + (1.625 - 0.5)\frac{F_2}{F_4} = 0.5 + 1.125 \cdot \frac{1}{3} = 0.875,$$

$$f(c_3) = 3.161,\ f(d_3) = f(c_2) = 2.85,$$

$$f(c_3) > f(d_3):$$

- Չորրորդ քայլում $a_4 = c_3 = 0.875,\ b_4 = b_3 = 1.625,\ c_4 = d_3 = 1.25,$

$$d_4 = 0.875 + (1.625 - 0.875)\frac{F_2}{F_3} = 1.25,$$

$$x_m = c_4 = d_4 = \frac{a_4 + b_4}{2} = 1.25:$$

Ի տարբերություն դիխոտոմիայի և ոսկե հատման մեթոդի, Ֆիբոնաչիի մեթոդում քայլերի քանակը տրվում է նախորոք (անհրաժեշտ չէ յուրաքանչյուր իտերացիայում ստուգել կանգառի պայմանը): Ֆիբոնաչիի մեթոդը օպտիմալ է ֆունկցիայի արժեքների քանակի հաշվման տեսակետից: Այսպես, մեր դիտարկած օրինակում միևնույն ճշտությունը ապահովելու համար դիխոտոմիայի մեթոդով անհրաժեշտ եղավ հաշվել ֆունկցիայի արժեքները վեց կետերում, ոսկե հատման դեպքում՝ հինգ, իսկ Ֆիբոնաչիի մեթոդով՝ ընդամենը չորս կետերում: Նկատենք նաև, որ մեծ n-ի դեպքում Ֆիբոնաչիի մեթոդը վերածվում է ոսկե հատման մեթոդի:

Առաջադրանք: Կիրառելով վերևում շարադրված չորս մեթոդները՝ պասիվ, դիխոտոմիայի, ոսկե հատման և ֆիբոնաչիի, կազմել ծրագիր որևէ մի ալգորիթմական լեզվով, որը կիրականացնի

$$f(x) = \begin{cases} -e^{3x/2},\ 0 \le x \le 0.45, \\ 2.3x - 3,\ 0.45 \le x \le 0.7, \\ x^2 - 1,\ 0.7 \le x \le 1 \end{cases}$$

ունիմոդալ ֆունկցիայի մինիմիզացիան $\varepsilon = 0.01$ ճշտու– թյամբ:

2.2 Առաջին և երկրորդ կարգի մեթոդներ

Շոշափողների մեթոդը: Դիցուք $f(x)$ ֆունկցիան ուռուցիկ է և դիֆերենցելի $[a, b]$ հատվածի վրա: Դիցուք $f'(a) < 0,\ f'(b) > 0$: Ընդունենք $a_1 = a,\ b_1 = b$: Գրաֆիկի $A(a, f(a))$ և $B(b, f(b))$ կետերից տանենք շոշափողներ և գտնենք նրանց հատման C կետի աբցիսը: Նշանակենք այն

c_1-ով: Եթե $f'(c_1) > 0$, ապա ընդունում ենք $a_2 = a_1,\ b_2 = c_1$: Եթե, $f'(c_1) < 0$, ապա ընդունում ենք $a_2 = c_1,\ b_2 = b_1$: Եթե $f'(c_1) = 0$, ապա համարում ենք, որ c_1-ը որոնելի կետն է: Այս պրոցեդուրան կրկնում ենք այնքան անգամ մինչև որ տեղի ունենան

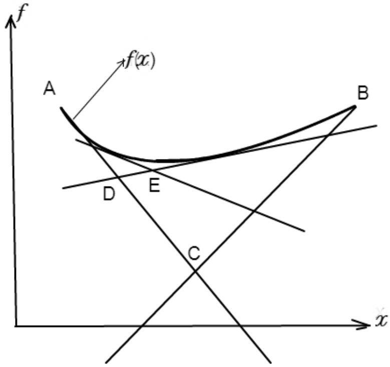
$$| b_k - a_k | \leq \varepsilon$$

կամ

$$| f'(c_k) | \leq \varepsilon$$

անհավասարությունները: Այս դեպքում c_k-ն համարում ենք որպես ֆունկցիայի մինիմումի կետ և ալգորիթմը ավարտվում է: Այս մեթոդը դասվում է առաջին կարգի մեթոդների շարքին, քանի որ ֆունկցիայի մինիմիզացիայի ընթացակարգում օգտագործվում են միայն նրա առաջին կարգի ածանցյալները: Հասկանալի է, որ այս մեթոդի իրագործումը օգտակար է այն դեպքերում, երբ ֆունկցիայի ածանցյալի հաշվումը կապված չէ որոշակի դժվարությունների հետ: Նշենք նաև, որ այս մեթոդի իրականացումը հենված է ուռուցիկ ֆունկցիայի կպոր առ կպոր գծային մոտարկման գաղափարի վրա: Իր պարզության պատճառով այս մեթոդը հաճախ կիրառվում է բավականաչափ ողորկ ունիմոդալ ֆունկցիաների մինիմիզացիայի ժամանակ:

Գծ.2.11-ում պրված է շոշափողների մեթոդի երկրաչափական մեկնաբանությունը:



Գծ. 2.11: Շոշափողների մեթոդի երկրաչափական մեկնաբանությունը

Նյուտոն-Ռաֆսոնի մեթոդը: Մեր նպատակն է գտնել $f'(x) = 0$ հավասարման արմատները:

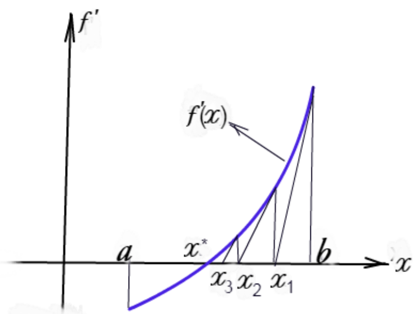
Սկզբնական x_0 կետից տանենք շոշափող f' ֆունկցիայի գրաֆիկին: Գտնենք այդ ուղղի և OX առանցքի հետ հատման կետը: Նշանակենք այն x_1-ով և այդ կետով նորից տանենք շոշափող f' ֆունկցիայի գրաֆիկին: Այս պրոցեդուրան կրկնում ենք այնքան անգամ մինչև, որ տեղի ունենա $| \ f'(x_k) \ | \leq \ \varepsilon$ անհավասարությունը, որտեղ ε-ը նախորոք տրված ճշտություն է: Այս դեպքում x_k-ն համարում ենք ֆունկցիայի մինիմումի կետ և ալգորիթմն ավարտվում է: $\{x_k\}$ հաջորդականությունը կառուցվում է հետևյալ կերպ: Գրում ենք x_k կետում f' ֆունկցիայի գրաֆիկին տարված շոշափող ուղղի

$$y = f'(x_k) + f''(x_k)(x - x_k)$$

հավասարումը և գտնում այս ուղղի և OX առանցքի հատման կետը, որը որոշվում է հետևյալ բանաձևով՝

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}:$$

Այս մեթոդը կոչվում է երկրորդ կարգի մեթոդ, որովհետև այստեղ օգտագործվում են ֆունկցիայի երկրորդ կարգի ածանցյալի արժեքները: Նյուտոն-Ռաֆսոնի մեթոդի գրաֆիկական մեկնաբանությունը տրված է գծ. 2.12-ում:



Գծ. 2.12: Նյուտոն-Ռաֆսոնի մեթոդի երկրաչափական մեկնաբանությունը

Օրինակ: Գտնել $f(x) = x + \frac{2}{x}$ ֆունկցիայի մինիմումի կետը $\varepsilon = 0.5$ ճշտությամբ, $x_0 = 0.5$: Քայլ առ քայլ նկարագրենք Նյուտոնի մեթոդը այս օրինակի դեպքում: Ունենք

$$f'(x) = 1 - \frac{2}{x^2}, \; f''(x) = \frac{4}{x^3}:$$

- Առաջին քայլում $f'(x_0) = -7$, $\mid f'(x_0) \mid > \varepsilon$, ուրեմն կանգառի պայմանը փեղի չունի: Ուսփի հաշվարկը շարունակում ենք.

$$f''(x_0) = 32, \ x_1 = 0.5 - \frac{-7}{32} = 0.719:$$

- Երկրորդ քայլում $f'(x_1) = -2.869$, $\mid f'(x_1) \mid > \varepsilon$: Ալգորիթըը շարունակվում է, քանի որ կանգառի պայմանը այս քայլում նույնպես չի կափարվում:

$$f''(x_1) = 10.762, \ x_2 = 0.719 - \frac{-2.869}{10.762} = 0.986:$$

- Երրորդ քայլում $f'(x_2) = -1.057$, $\mid f'(x_2) \mid > \varepsilon$, հեփևաբար, հաշվարկը շարունակում ենք.

$$f''(x_2) = 4.173, x_3 = 0.986 - \frac{-1,057}{4.173} = 1.239 :$$

- Չորրորդ քայլում $f'(x_3) = -0.303$, $\mid f'(x_3) \mid < \varepsilon$: Ինչպես փեսնում ենք այս քայլում կանգառի պայմանը արդեն փեղի ունի, ուսփի ալգորիթմը ավարփվում է և ընդունում ենք

$$x_m = x_3 = 1.239:$$

Վերջում նշենք, որ ֆունկցիայի մինիմիզացիայի պրոցեսում որքան շափ ինֆորմացիա իմանանք ֆունկցիայի մասին, այնքան արագ կհասնենք անհրաժեշփ արդյունքին, այսինքն՝ մինիմիզացիայի պրոցեսի արագությունը մեծանում է: Սակայն դրան զուգահեռ շափանում է նաև անհրաժեշփ հաշվարկների քանակը: Այսպես, վերևում դիփարկված մեթոդներից բոլորից դանդաղ զուգամիտում են զրոյական կարգի մեթոդները, հետո՝ առաջին կարգինը, այնուհետև՝ երկրորդ կարգի մեթոդները: Բայց, օրինակ, շոշափողների մեթոդում, ի տարբերություն ոսկե հատման մեթոդի, անհրաժեշտ է հաշվել ոչ միայն ֆունկցիայի արժեքները, այլև՝ նրա ածանցյալի: Իսկ Նյուտոնի մեթոդում լրացուցիչ պետք է հաշվել նաև ֆունկցիայի երկրորդ կարգի ածանցյալի արժեքները:

Հարկ ենք գտնում նաև ասել, որ մեկ փոփոխականի ֆունկցիայի մինիմիզացիայի ընթացակարգերի վերաբերյալ ավելի մանրամասն ու ամբողջական շարադրանք կարելի է գտնել [7, 24] աշխատանքներում:

Գլուխ 3

Բազմաչափ օպտիմիզացիա: Գրադիենտային մեթոդները

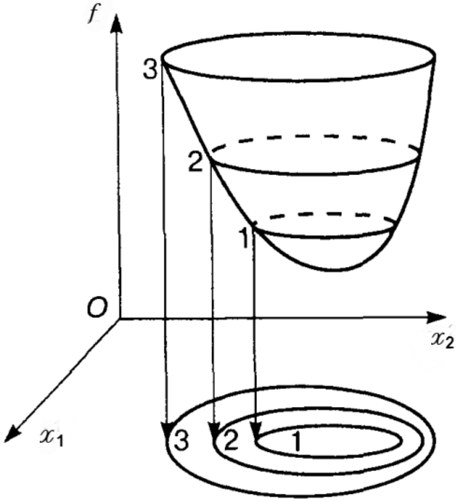
Սահմանում 3.0.1: $f(x)$ *Ֆունկցիայի մակարդակի գիծ կոչվում է*

$$V_C = \{x \in R^n / f(x) = C\}$$

բազմությունը, որտեղ $C-$*ն կամայական թիվ է (տես գծ.3.1):*

Ֆունկցիայի մակարդակի գծերին անվանում են նաև մակարդակի բազմություններ:

Սովորաբար ֆունկցիայի մակարդակի գծերը փակ կորեր են: Դրանց միջոցով հնարավոր է լինում գրաֆիկորեն մեկնաբանել ֆունկցիայի մինիմիզացիայի ալգորիթմները, որոնք կշարադրվեն այս գլխում: Որոշ դասի ֆունկցիաների համար կկառուցվեն այդ կորերը: Ընդհանուր դեպքում նրանց գտնելու հարցը բարդ է:



Գծ. 3.1: Ֆունկցիայի մակարդակի կորերը $C = 1, C = 2, C = 3$ արժեքների դեպքում

Օրինակ: Կառուցել $f(x) \equiv f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$ ֆունկցիայի մակարդակի կորերը: f-ը ներկայացնենք հետևյալ տեսքով.

$$f(x_1, x_2) = \frac{1}{2}(Ax, x):$$

Այս դեպքում

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}:$$

Այնուհետև A մատրիցը բերում ենք կանոնական տեսքի: Դրա համար գտնենք նրա սեփական արժեքները.

$$det(A - \lambda E) = 0:$$

Այսպեղից $\lambda_1 = 3,\ \lambda_2 = 1$: Հետևաբար այդ արժեքներին համապատասխանող սեփական միավոր վեկտորներն են՝

$$e_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}),\ e_2 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}):$$

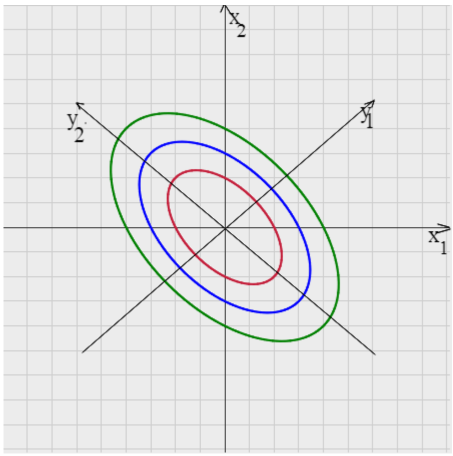
Նշանակենք՝

$$\begin{cases} y_1 = \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2, \\ y_2 = -\frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2 : \end{cases}$$

Այսպիսով, կոորդինապային նոր $\{y_1, y_2\}$ համակարգում f ֆունկցիան ներկայացվում է հեպևյալ պեսքով.

$$f(y) = \frac{3}{2}y_1^2 + \frac{1}{2}y_2^2:$$

Այսինքն՝ f ֆունկցիայի մակարդակի գծերը էլիպսներ են, որոնք ներկայացված են գծ.3.2-ում:



Գծ. 3.2: $x_1^2 + x_1x_2 + x_1^2$ ֆունկցիայի մակարդակի կորերը

3.1 Կոորդինատային իջեցման մեթոդը

Այս մեթոդով $f(x_1, x_2, \dots, x_n)$ ֆունկցիայի մինիմիզացիան իրագործվում է n քայլանի ցիկլերով:

Դիցուք
$$x^k = (x_1^k, \dots, x_n^k):$$

- Ամրագրում ենք $x_2^k, x_3^k, \dots, x_n^k$ արժեքները և կատարում ենք մինիմիզացիա ըստ x_1 փոփոխականի՝ ստանալով $(x^{k+1}, x_2^k, \dots, x_n^k)$ կետը, և դրանով իսկ առաջին քայլը ավարտվում է:

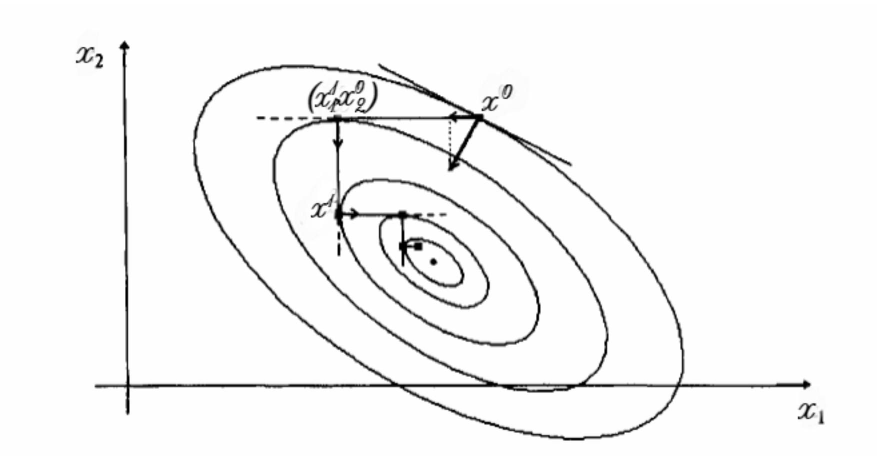
- Ամրագրում ենք $x_1^{k+1}, x_3^k, \dots, x_n^k$ արժեքները և կատարում մինիմիզացիա ըստ x_2 փոփոխականի՝ ստանալով $(x_1^{k+1}, x_2^{k+1}, \dots, x_n^k)$ կետը:

- $\dots$

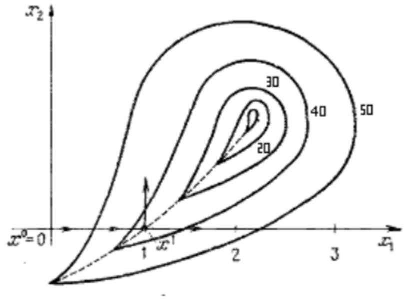
- n-րդ քայլում ամրագրում ենք $x_1^{k+1}, x_2^{k+1}, \dots, x_{n-1}^{k+1}$ կետեր և կատարում ֆունկցիայի մինիմիզացիա ըստ x_n փոփոխականի: Արդյունքում ստանում ենք $(x_1^{k+1}, \dots, x_n^{k+1})$ կետը:

Այսպիսով, n քայլերից հետո առաջին ցիկլն ավարտվում է ստանալով՝ $x^{k+1} = (x_1^{k+1}, \dots, x_n^{k+1})$ կետը: Պրոցեսը շարունակվում է մինչև որ կատարվի կանգառի քայլը.
$$| f(x^{k+1}) - f(x^k) | \le \varepsilon \text{ կամ } \|x^{k+1} - x^k\| \le \varepsilon:$$

Կոորդինատային իջեցման մեթոդը գրաֆիկորեն պատկերված է գծ. 3.3-ում: Նշենք, որ կան ֆունկցիաներ, որոնց համար կոորդինատային մեթոդը չի բերում մինիմումի կետի: Դա հաճախ տեղի է ունենում այն դեպքերում, երբ ֆունկցիայի մակարդակի բազմությունները ունենում են, այսպես կոչված, կիրճային փեսք, այսինքն՝ երբ նրանք խիստ ձգված են ինչ որ մի ուղղությամբ: Օրինակ՝ գծ. 3.4-ի դեպքում x^1 կետից կոորդինատային x_2 առանցքին զուգահեռ շարժումը չի բերում իջեցման՝ այդ ուղղությամբ կամայական շարժում բերում է ֆունկցիայի արժեքների աճին: Այսպիսի դեպքերում ասում են, որ իջեցման պրոցեսը արգելակվում է:



Գծ. 3.3: Կոորդինատային իջեցման մեթոդի երկրաչափական մեկնաբանությունը. կոորդինատային առանցքներին զուգահեռ կատարված շարժման հետագծերը շոշափում են ֆունկցիայի մակարդակի գծերը:



Գծ. 3.4: Կոորդինատային իջեցման պրոցեսի արգելակում

Օրինակ: Գտնել $f(x) = 9x_1^2 + x_2^2$ ֆունկցիայի մինիմումի կետը $\varepsilon = 0.5$, ճշտությամբ: Որպես մեկնարկային արժեք վերցնենք $x^0 = (1, 1)$ կետը: Իսկ կանգառի քայլ համարենք

$$\|x^{k+1} - x^k\| \leq \varepsilon$$

պայմանը:

- Ամրագրենք $x_2 = x_2^0 = 1$ և կատարենք ֆունկցիայի մինիմիզացիա ըստ x_1 փոփոխականի: Այսինքն՝ պետք է մինիմիզացնել $f(1+\alpha, 1) = 9(1+\alpha)^2 + 1$ ֆունկցիան $(-\infty, +\infty)$ միջակայքի վրա: Ակնհայտ է, որ $\alpha = -1$-ը մինիմումի կետն է և հետևաբար $x_1^1 = 0$:

- Ամրագրենք $x_1 = x_1^1 = 0$ և կատարենք մինիմիզացիա ըստ x_2 փոփոխականի: Այսինքն՝ գտնենք $f(0, 1+\alpha) = (1 + \alpha)^2$ ֆունկցիայի մինիմումի կետը $(-\infty, +\infty)$ բազմության վրա: Ակնհայտ է, որ $\alpha = -1$-ը մինիմումի կետն է: Վերջնականորեն առաջին ցիկլից հետո ստանում ենք $x^1 = (0, 0)$: Բայց կանգառի պայմանը տեղի չունի, քանի որ

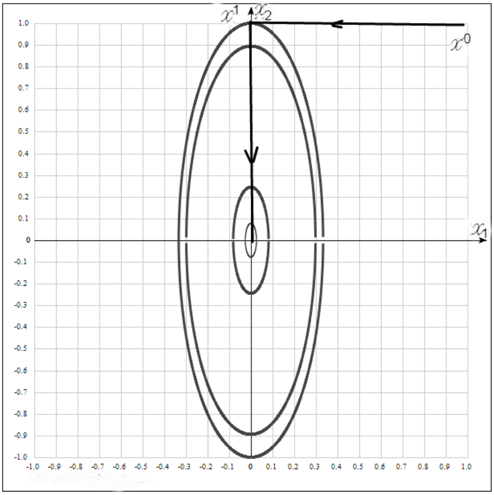
$$\mid x^1 - x^0 \mid = \sqrt{2} > 0.5:$$

Ուսփի հաշվարկը շարունակում ենք:

- Նախորդին անալոգ սպանում ենք՝

$$x^2 = (0,0):$$

Մյուս կողմից, քանի որ $| \ x^2 - x^1 \ | = 0 < 0.5$, ապա ալգորիթմը ավարտվում է և ընդունում ենք, որ $x_m = x^2 = (0,0)$-ը ֆունկցիայի մինիմումի կետն է, որը փվյալ դեպքում ճշգրիտ մինիմումն է (տե՛ս գծ.3.5):



Գծ. 3.5: Կոորդինապային իջեցման մեթոդը ($f(x_1, x_2) = 9x_1^2 + x_2^2$ ֆունկցիայի դեպքը)

ԽՆԴԻՐՆԵՐ

1. Կառուցել հետևյալ ֆունկցիաների մակարդակի գծերը.

   ա) $f(x_1, x_2) = x_1^2 + x_2^2$,

   բ) $f(x_1, x_2) = x_2^2 - x_1^2$,

   գ) $f(x_1, x_2) = x_1^2 - x_1x_2 + x_2^2$:

   Կոորդինատային իջեցման մեթոդով լուծել հետևյալ խնդիրները և գրաֆիկորեն ցույց տալ իջեցման պրոցեսի շարժման հետագծերը.

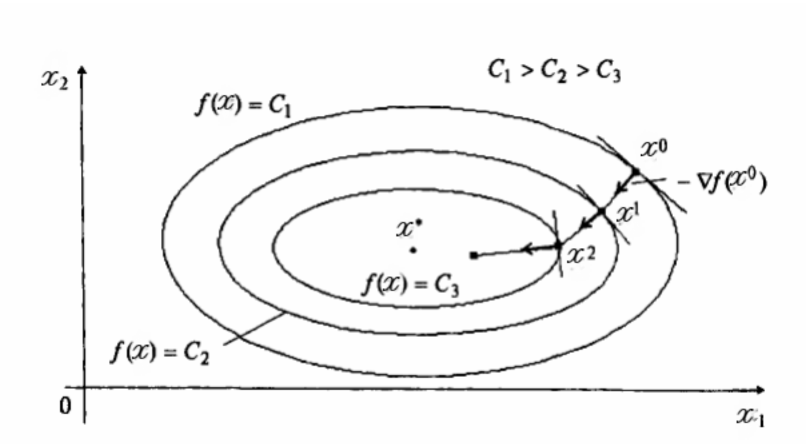
2. $f(x) = 2x_1^2 + x_2^2 \to min$:

3. $f(x) = 2x_1^2 + x_2^2 + x_1x_2 \to min$:

3.2 Գրադիենտային մեթոդը

Գրադիենտային մեթոդը դասվում է դիֆերենցելի ֆունկցիաների մինիմիզացիայի թվային հիմնական մեթոդների շարքին: Այդ մեթոդի էությունը շատ պարզ է: Եթե $f(x)$ ֆունկցիան դիֆերենցելի է, ապա նրա հակագրադիենտը՝ $h = -f'(x)$ վեկտորը, յուրաքանչյուր x կետում ցույց է տալիս ֆունկցիայի **նվազման ուղղությունը:** Վերցվում է սկզբնական կետ և կատարվում ոչ մեծ քայլ հակագրադիենտի ուղղությամբ: Արդյունքում ստացվում է նոր կետ, որում ֆունկցիայի արժեքը փոքր է սկզբնականից: Այս կետում պրոցեդուրան կրկնվում է: Այս պրոցեսի ընթացքում շարժում է կատարվում ֆունկցիայի նվազման ուղղությամբ: Դա հիմք է տալիս ենթադրել, որ կամայական սկզբնական x^0 կետից սկսվող և $x^{k+1} = x^k - \alpha_k f'(x_k)$ $(\alpha_k > 0)$ ռեկուրենտ առնչությամբ կառուցված հաջորդականության անդամները մեծ k ինդեքսների դեպքում մոտ կլինեն f-ի մինիմումի կետին (տե՛ս գծ.3.6):

Պետք է նշել, որ մեծ քայլերի ընտրության դեպքում հնարավոր է պրոցեսը լինի փարամետ, իսկ փոքր քայլերի դեպքում զուգամիտության պրոցեսը կարող է երկարել:



Գծ. 3.6: Գրադիենտային վայրէջքի պրոցես

Այս գլխում գրադիենտային մեթոդի զուգամիտությունը հիմնավորվում է որոշ դասի ֆունկցիաների և α_k թվերի հատուկ եղանակներով ընտրությունների դեպքերում: Ֆունկցիայի մինիմիզացիայի բազմազան ալգորիթմների և նրանց պրակտիկ իրականացումների հետ ավելի մանրամասն կարելի է ծանոթանալ [1,5,10,20,29] աշխատանքներում:

3.3 Ֆունկցիայի մինիմիզացիայի գրադիենտային եղանակների ընդհանուր նկարագրությունը

Սահմանում 3.3.1: *h վեկտորը կոչվում է f ֆունկցիայի նվազման ուղղություն x կետում, եթե բավականաչափ փոքր դրական α թվերի համար տեղի ունի*

$$f(x + \alpha h) < f(x)$$

անհավասարությունը:

Այլ խոսքով, h-ը այն վեկտորն է, որի ուղղությամբ բավականաչափ փոքր շարժվելիս կարելի է ֆունկցիայի արժեքները փոքրացնել:

Լեմմ 3.3.1: *Դիցուք f-ը դիֆերենցելի է x կետում և h-ն այնպիսի վեկտոր է, որ*

$$(f'(x), h) < 0:$$

Այդ դեպքում h-ը f-ի նվազման ուղղություն է x կետում:

► Քանի որ f-ը դիֆերենցելի է, ապա

$$f(x + \alpha h) = f(x) + \alpha[(f'(x), h) + \frac{o(\alpha)}{\alpha}]:$$

Այստեղից բավականաչափ փոքր դրական α թվերի համար միջակ փակագծերում գրած արտահայտությունը դառնում է բացասական: Այնպես որ կստանանք

$$f(x + \alpha h) < f(x): \blacksquare$$

Դիտողություն: Մասնավորապես, որպես նվազման ուղղություն կարելի է վերցնել $h = -f'(x)$, որը կոչվում է հակագրադիենտի ուղղություն: Կարելի է ցույց տալ, որ հակագրադիենտը ցույց է տալիս է ֆունկցիայի ամենաարագ նվազման ուղղությունը:

Գրադիենտային վայրէջքի մեթոդները կառուցվում են հետևյալ կերպ: Ընտրվում է կամայական x^0 կետ R^n տարածությունից և կառուցվում է

$$x^{k+1} = x^k + \alpha_k h^k,\ k = 0, 1, 2, ... \qquad (3.1.1)$$

ռեկուրենտ առնչությունը: Այստեղ $h^k = -f'(x^k)$, իսկ α_k թվերը կոչվում են քայլեր, որոնք ընտրվում են որոշակի օրինաչափությամբ:

Օրինակ: Դիտարկենք $f(x) = ax^2$ ֆունկցիայի համար գրադիենտային մեթոդը.

$$x_{k+1} = x_k(1 - 2\alpha_k a):$$

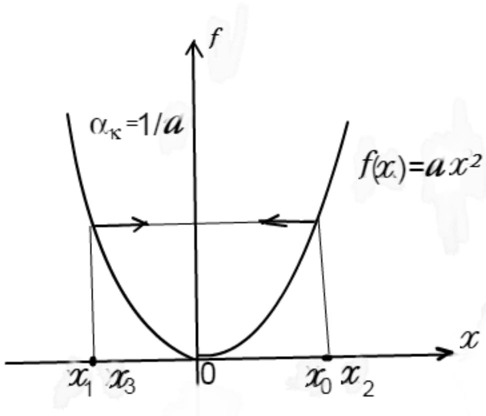
Ակնհայտ է, որ այս պրոցեսը զուգամետ կլինի զրոյի, եթե $\alpha_k < \frac{1}{a}$: $\alpha_k = \frac{1}{a}$-ի դեպքում պրոցեսը ցիկլիկ է՝

$$x_1 = -x_0,\ x_2 = x_0,\ x_3 = -x_0, \ldots (\text{տես գծ.3.7}):$$

Իսկ $\alpha_k > \frac{1}{a}$-ի դեպքում պրոցեսը տարամետ է:

Ինչպես երևում է այս պարզ օրինակից, գրադիենտային մեթոդի զուգամիտությունը կախված է α_k քայլի երկարության ընտրությունից: Այս պարագրաֆում տրվում են քայլի ընտրության մի քանի եղանակների ընդհանուր նկարագրություններ: Այդ եղանակներից որոշների զուգամիտության և զուգամիտության արագության հարցեր դիտարկվում են առանձին պարագրաֆում: Հայտնի են քայլի ընտրության մի քանի եղանակներ, որոնք ներկայացվում են ստորև:

1. **Քայլի ընտրության ապրիորի եղանակը:** Այս դեպքում α_k քայլերը դրական թվեր են, որոնք բավարարում են



Գծ. 3.7: Ցիկլիկ գրադիենտային պրոցես

հետևյալ պայմաններին.

$$\sum \alpha_k = +\infty, \; \sum \alpha_k^2 < +\infty:$$

Մասնավորապես, այս պայմաններին բավարարում են

$$\alpha_k = \frac{1}{k+1}, \;\; \alpha_k = \frac{ln(k+1)}{k+1}, \; k = 0, 1, 2, ...$$

թվային հաջորդականությունները:

2. **Քայլի ընտրության ամենաարագ վայրէջքի եղանակը:** Այս դեպքում $\alpha_k > 0$ քայլերը ընտրվում են հետևյալ կերպ: Կազմում են $g(\alpha) \equiv f(x^k + \alpha h^k)$ ֆունկցիան և գտնում նրա մինիմումի կետը $(0, \infty)$ միջակայքի վրա: Այն կետը, որտեղ մինիմումը հասանելի է համարվում է α_k, այսինքն՝

$$\alpha_k \equiv arg \min_{\alpha > 0} g(\alpha):$$

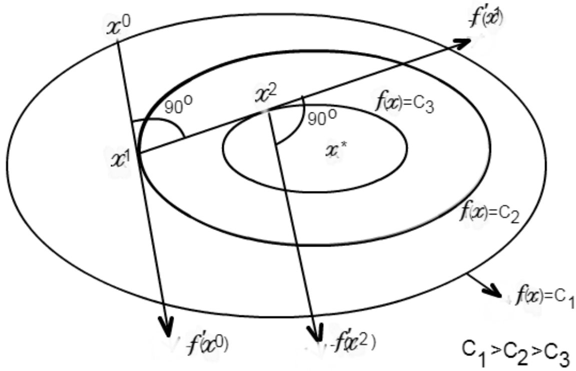
Նշենք, որ այս մեթոդում շարժման ուղղությունները երկու հաջորդական իտերացիաներում իրար փոխուղղահաց են:

Իրոք, քանի որ α_k քայլը ընտրում ենք $g(\alpha) \equiv f(x^k + \alpha h^k)$ ֆունկցիայի մինիմիզացիայի պայմանից, ապա

$$g'(\alpha_k) = -(f'(x^{k+1}), f'(x^k)) = 0:$$

Այսպիսով, x^k կետից շարժման ուղղությունը շոշափում է մակարդակի բազմությանը x^{k+1} կետում (տես գծ.3.8): Այս մեթոդի առավելությունը այն է, որ եթե k-րդ քայլում մինիմումի կետը ընկած լինի $x^k + \alpha h^k$ ճառագայթի վրա, ապա ամենաարագ վայրէջքով այդ քայլում ճշգրիտ կգտնենք ֆունկցիայի մինիմումի կետը: Հետագայում մենք կտեսնենք, որ այս պարզ փաստը կարևոր նշանակություն ունի համալուծ գրադիենտների մեթոդում քառակուսային ֆունկցիայի մինիմիզացիայի համար վերջավոր ալգորիթմի մշակման ժամանակ: Իսկ այս մեթոդի թերությունը նրանում է, որ յուրաքանչյուր քայլում պետք է կատարել միաչափ մինիմիզացիա անվերջ միջակայքի վրա, ինչը պրակտիկ տեսակետից հեշտ իրագործելի խնդիր չէ: Այդ խնդիրը, սովորաբար, լուծվում է մոտավոր՝ կիրառելով թվային որոշ մեթոդներ: Բայց, այնուամենայնիվ, պրակտիկան ցույց է տալիս, որ այս մեթոդը պահանջում է ավելի քիչ գործողություններ, քան մնացած գրադիենտային եղանակները և նրա զուգամիտության արագությունը մեծ է:

Պարզագույն դեպքերում հնարավոր է լինում գտնել α_k մեծությունը բացահայտ տեսքով:



Գծ. 3.8: Ամենաարագ վայրէջքի մեթոդի երկրաչափական մեկնաբանությունը. շարժման ուղղություններն երկու հաջորդական իտերացիաներում իրար փոխուղղահայաց են:

Օրինակ: Ամենաարագ վայրէջքով գտնել $f(x) = 9x_1^2 + x_2^2$ ֆունկցիայի մինիմումի կետը $\varepsilon = 0.05$ ճշտությամբ, որպես սկզբնական մոտավորություն վերցնել $x^0 = (1,1)$ կետը: Կանգառի քայլը վերցնել $\|f'(x^k)\| \le \varepsilon$ պայմանը: Հաշվարկը կատարում ենք ստորակետից հետո երեք նիշի ճշտությամբ: Ունենք $f'(x_1, x_2) = (18x_1, 2x_2)$: Հետևաբար

- առաջին քայլում ունենք $x^1 = x^0 - \alpha f'(x^0) = (1,1) - \alpha(18,2)$:

  Այստեղից

  $$f(x^0 - \alpha f'(x^0)) = 9(1-18\alpha)^2 + (1-2\alpha)^2:$$

  Ուստի

  $$\alpha_0 = argmin_{\alpha>0}(f(x^0 - \alpha f'(x^0)) = 0.056,$$

$$x^1 = x^0 - \alpha_0 f'(x^0) = (-0.008, 0.888):$$

- Երկրորդ քայլում $f'(x^1) = (-0.144, 1.776)$: Ստուգենք կանգառի քայլը.

$$\|f'(x^1)\| = \sqrt{(\frac{\partial f(x^1)}{\partial x_1})^2 + (\frac{\partial f(x^2)}{\partial x_2})^2} < \varepsilon :$$

Այս անհավասարությունը տեղի չունի, հետևաբար հաշվակը շարունակում ենք.

$$\alpha_1 = 0.475,\ x^2 = (0.06, 0.044):$$

- Երրորդ քայլում ունենք՝

$$f'(x^2) = (1.087, 0.089),\ \|f'(x^2)\| > \varepsilon, \alpha_2 = 0.056,$$

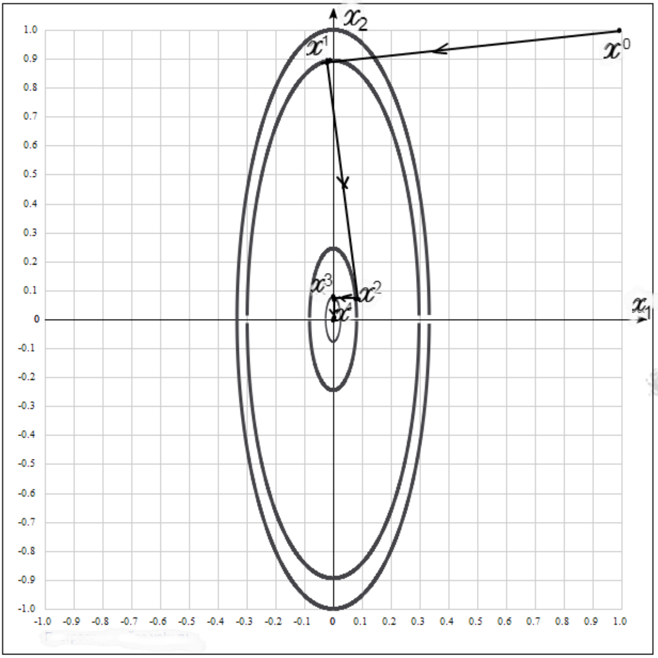
$$x^3 = (0.000, 0.039):$$

- Չորրորդ քայլում կստանանք՝

$$f'(x^3) = (0.000, 0.078),\ \|f'(x^3)\| > \varepsilon,\ \alpha_3 = 0.456,$$

$$x^4 = (0.000, 0.003):$$

- Հինգերորդ քայլում կանգառի $\|f'(x^4)\| < \varepsilon$ պայմանը կատարվում է, հետևաբար ընդունում ենք $x_m = x^4$ և ալգորիթմն ավարտվում է: Գծ. 3.9-ում պատկերված է շարժման հետագիծը: Ինչպես երևում է գծագրից երկու մեծ քայլերից հետո հետագիծը հայտնվում է կիրճի հատակում: Այնուհետև բավականաչափ փոքր քայլերով այն մոտենում է մինիմումի կետին:



Գծ. 3.9: Ամենաարագ վայրէջքի մեթոդը ($f(x_1, x_2) = 9x_1^2 + x_2^2$ ֆունկցիայի դեպքը)

Որոշ դասի ֆունկցիաների համար կարելի է ստանալ անալիտիկ բանաձևեր α_k քայլերի որոշման համար, ինչպես հետևյալ օրինակում:

Օրինակ: Դիցուք $f(x) = 1/2(\mathbf{A}x, x) + (b, x)$, որտեղ $\mathbf{A}$-ն $(n \times n)$ չափանի սիմետրիկ, դրական որոշյալ մատրից է, իսկ b-ն n չափանի վեկտոր է: Կարելի է ցույց տալ, որ f-ը ուռուցիկ է, ունի միակ մինիմումի կետ R^n-ի վրա և նրա գրադիենտը որոշվում է հետևյալ բանաձևով. $f'(x) = \mathbf{A}x + b$: Այս դեպքում ունենք

$$g(\alpha) \equiv f(x^k + \alpha h^k) = 1/2(\mathbf{A}(x^k + \alpha h^k), x^k + \alpha h^k)) +$$

$$+(b, x^k + \alpha h^k) = 1/2(\mathbf{A}h^k, h^k)\alpha^2 +$$

$$+(\mathbf{A}x^k + b, h^k)\alpha + (1/2\mathbf{A}x^k + b, x^k):$$

Դժվար չէ նկատել, որ ստացված քառակուսային եռանդամը հասնում է իր փոքրագույն արժեքին R^n-ի վրա

$$\alpha_k = -\frac{(\mathbf{A}x^k + b, h^k)}{(\mathbf{A}h^k, h^k)} = -\frac{(f'(x^k), h^k)}{(\mathbf{A}h^k, h^k)} \geq 0 \qquad (3.1.2)$$

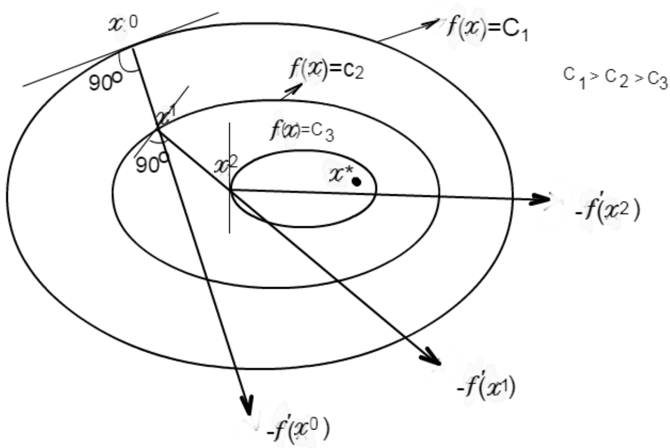
կետում: Հետևաբար

$$f(x^k + \alpha_k h^k) = \min_{\alpha \geq 0} f(x^k + \alpha h^k) = \min_{\alpha \in R^n} f(x^k + \alpha h^k):$$

3. **Քայլի ընտրության կիսման եղանակը:** Ֆիքսում ենք որևէ դրական ε թիվ $(0, 1)$ միջակայքից և $\alpha = 1$ թվի համար ստուգում ենք

$$f(x^k - \alpha f'(x^k)) - f(x^k) \leq -\varepsilon\alpha \|f'(x^k)\|^2 \qquad (3.1.3)$$

անհավասարությունը: Եթե այն տեղի ունի, ապա α_k-ն համարում ենք հավասար α-ի: Հակառակ դեպքում α-ն կիսում ենք և նորից ստուգում (3.1.3) անհավասարությունը և այսպես շարունակ: Եթե որևէ p-րդ քայլում բավարարվում է (2.1.3) անհավասարությունը, ապա $\alpha_k = \frac{1}{2^p}$: Այս մեթոդով շարժումը կատարվում է, որոշակի զիգզագաձև $x^0 x^1 x^2 \dots$ հետագծով, ընդ որում՝ յուրաքանչյուր x^k կետում հետագիծը ուղղահայաց է համապատասխան մակարդակի բազմությանը (տես գծ.3.10):



Գծ. 3.10: Քայլի կիսման մեթոդի երկրաչափական մեկնաբանությունը. $\{x^k\}$-ում շարժման հետագիծն ուղղահայաց է համապատասխան մակարդակի բազմությանը:

ԽՆԴԻՐՆԵՐ

---

1. Գտնել հետևյալ ֆունկցիաների էքստրեմումի կետերը ամենաարագ վայրէջքի մեթոդով.
   սկզբնական x^0 կետից սկսած կառուցել $\{x^k\}$ հաջորդականությունը (3.1.1) ռեկուրենտ առնչությամբ և քայլերի ընտրությունը կատարել (3.1.2) բանաձևով: Եթե $\|f'(x^k)\| < \varepsilon_0$, ապա պրոցեսն ավարտել և x^k-ն համարել ֆունկցիայի էքստրեմումի կետ:

Այստեղ $\varepsilon_0 > 0$ թիվը նախապես տրված ճշտությունն է:

ա) $f(x_1, x_2) = 5x_1^2 + 5x_2^2 - 6x_1x_2 \to min,\ x^0 = (1, 1),\ \varepsilon_0 = 0.01$,

բ) $-3x_1^2 - 2x_2^2 + x_1 + 6x_2 - 15 \to max,\ x^0 = (0, 1),\ \varepsilon_0 = 0.1$,

գ) $3x_1^2 + x_2^2 + 2x_1x_2 + x_2 \to min,\ x^0 = (1, 0),\ \varepsilon_0 = 0.01$:

2. Դիցուք f-ը երկու անգամ դիֆերենցելի է x կետում և

$$(f'(x), h) = 0,\ (f''(x)h, h) < 0:$$

Ապացուցել, որ h-ը f ֆունկցիայի նվազման ուղղություն է:

3. Դիցուք $H(x) = f''(x)$ հեսիանը դրական որոշյալ է և $f'(x) \neq 0$: Ապացուցել, որ

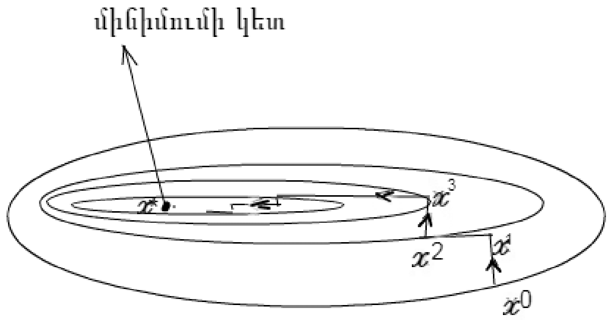
$$h = -H(x)^{-1} f'(x)$$

վեկտորը f ֆունկցիայի նվազման ուղղություն է:

Առաջադրանք: Գրել ծրագիր, որը մինիմիզացնում է $f(x) = 1/2(\mathbf{A}x, x) + (b, x)$ ֆունկցիան R^n-ի վրա: Այստեղ $\mathbf{A}$-ն $(n \times n)$ չափանի սիմետրիկ, դրական որոշյալ մատրից է, իսկ b-ն n չափանի վեկտոր է: Մուտքի տվյալներն են $\mathbf{A}, b, x^0, n, \varepsilon, \varepsilon_0$ պարամետրերը: Այստեղ ε_0-ն ճշտությունն է: Կառուցել $\{x^k\}$ հաջորդականությունը գրադիենտային իջեցման երեք մեթոդներով և համեմատել դրանք քայլերի քանակի տեսակետից: Եթե $\|f'(x^k)\| < \varepsilon_0$, ապա պրոցեսն ավարտել և համարել x^k-ն մինիմումի կետ: Համեմատել ստացված արդյունքները $x^* = -\mathbf{A}^{-1}b$ վեկտորի հետ, որը f-ի մինիմումի կետն է :

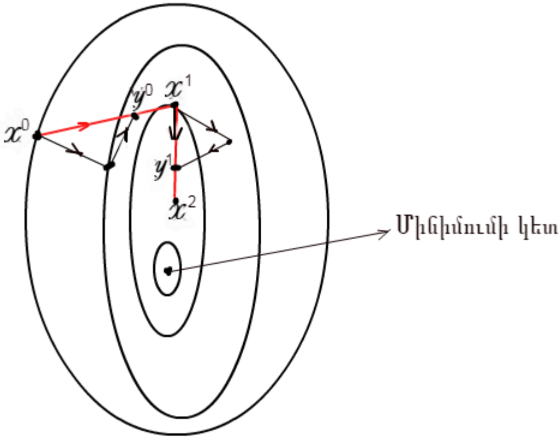
3.4 Արագացված գրադիենտային մեթոդները

Գրադիենտային մեթոդները զուգամիտում են շատ դանդաղ, երբ ֆունկցիայի մակարդակի բազմությունը խիստ ձգված է մի ուղղությամբ: Այս դեպքում գնալով քայլի երկարությունը փոքրանում է և զուգամիտության պրոցեսը երկարում է (տես գծ.3.11):



Գծ. 3.11: Գրադիենտային մեթոդի շարժման հետագիծը, երբ ֆունկցիայի մակարդակի բազմությունը նման է կիրճի:

Նման դեպքերում կիրառվում են հատուկ միջոցներ, որոնք հնարավորություն են փալիս կատարել մեծ քայլ մինիմումի կետի ուղղությամբ և դրանով իսկ արագ մոտենալ նրան: Այդ միջոցներից են հետևյալ երկու մեթոդները:



Գծ. 3.12: p-րդ կարգի արագացված գրադիենտային մեթոդի երկրաչափական մեկնաբանությունը

Առաջինը կոչվում է p-**րդ կարգի արագացված գրադիենտային մեթոդ:** Այս մեթոդում x^k կետից կատարում ենք p քայլ ամենաարագ վայրէջքի մեթոդով՝ հասնելով y^k կետ: x^{k+1} կետը ստանալու համար կատարում ենք միաչափ մինիմիզացիա $y^k - x^k$ վեկտորի ուղղությամբ x^k կետից: Սովորաբար վերցնում են $p = n$, որտեղ n-ը տարածության չափողականությունն է: Այսպեղ յուրաքանչյուր իտերացիայում պետք է լուծել $n + 1$ հատ միաչափ օպտիմիզացիայի խնդիրներ (տես գծ.3.12): Այսինքն, այս մեթոդը ավելի «թանկարժեք» է, քան ամենաարագ վայրէջքի մեթոդը: Այս մեթոդը մեկնաբանենք հետևյալ օրինակի միջոցով:

Օրինակ: Գտնել $f(x) = 9x_1^2 + x_2^2$ ֆունկցիայի մինի–մումի կետը $\varepsilon = 0.05$ ճշտությամբ: Որպես սկզբնական մոտավորություն վերցնել $x^0 = (1, 1)$, իսկ կանգառի պայման համարել $\|f'(x^k)\| \leq \varepsilon$ անհավասարությունը:

Սկզբնական $x^0 = (1, 1)$ կետից ամենաարագ վայրէջքի մեթոդով կատարենք երկու քայլ: Արդյունքում կստանանք $y^0 = (0.06, 0.044)$: Այնուհետև, գտնում ենք f ֆունկցիայի մինիմումի կետը $x^0 + \alpha(y^0 - x^0)$ ճառագայթի վրա: Այսինքն՝ կատարում ենք մեկ փոփոխականի $\varphi(\alpha) \equiv f(x^0 + \alpha(y^0 - x^0))$ ֆունկցիայի մինիմիզացիա $[0, \infty)$ միջակայքի վրա: Ունենք

$$f(x^0 + \alpha(y^0 - x^0)) = 9(1 - 0.94\alpha)^2 + (1 - 0.956\alpha)^2:$$

Այստեղից

$$argmin_\alpha \varphi(\alpha) = 1.062,$$

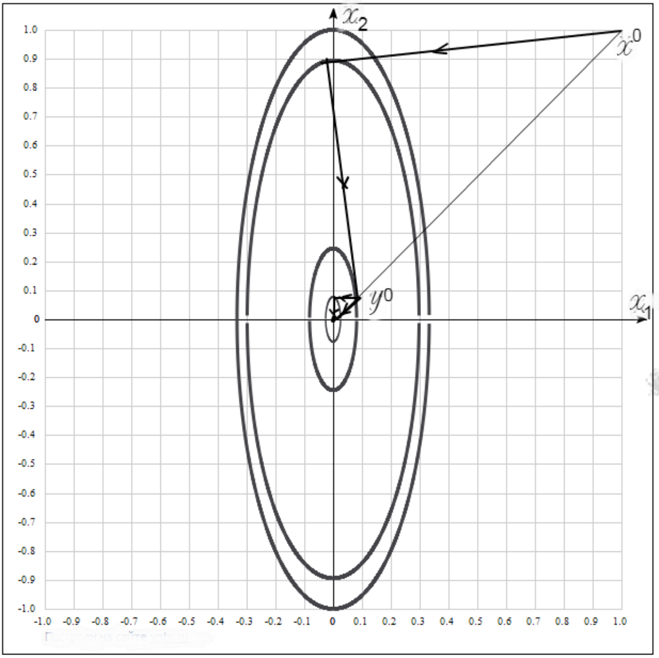
և հետևաբար

$$x^1 = (1 - 0.94.1.062, 1 - 0956.1.062) = (0.002, -0.015):$$

Քանի որ $\|f'(x^1)\| = 0.043 \leq \varepsilon$, ապա

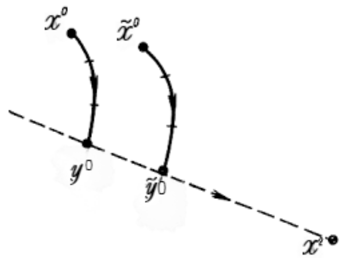
$$x_m = (0.002, -0.015):$$

Նկատենք, որ այս մեթոդով կատարվում է միաչափ մինիմիզացիայի մեկ քայլ պակաս, քան ամենաարագ վայրէջքում: Գծ. 3.13-ում պատկերված են ամենաարագ և արագացված գրադիենտային մեթոդների հետագծերը դիտարկված օրինակի համար: f Ֆունկցիայի մակարդակի բազմությունները էլիպսներ են ձգված OX_2 առանցքով: Ինչպես երևում է գծագրից, եթե կիրառում ենք ամենաարագ վայրէջքի մեթոդը, ապա երկու քայլից հետո իջնում ենք կիրճի հատակ և դրանից հետո փոքր քայլերով զիգզագաձև ճանապարհով մոտենում ենք մինիմումի կետին: Իսկ արագացված գրադիենտային մեթոդի դեպքում, երկրորդ քայլից հետո մեծ քայլ է կատարվում մինիմումի կետի ուղղությամբ:



Գծ. 3.13: Ամենաարագ վայրէջքի և արագացված գրադիենտային մեթոդների հետագծերը ($f(x_1, x_2) = 9x_1^2 + x_2^2$ ֆունկցիայի դեպքը)

Երկրորդ եղանակը, որն արագացնում է գրադիենտային մեթոդը կրում է Գելֆանդի անունը և այն կոչվում է **կիրճային** եղանակ: Այստեղ վերցնում ենք սկզբնական x^0 մոտավորությանը մոտիկ $\tilde{x}^0$ կետը և այս երկու կետերից կատարում ենք գրադիենտային իջեցում` ստանալով $y^0$, $\tilde{y}^0$ կետերը: Այնուհետև կատարվում է միաչափ մինիմիզացիա $\tilde{y}^0 - y^0$ վեկտորի ուղղությամբ $y^0$ սկզբնակետով` ստանալով x^1 կետը: Դրանից հետո x^1 կետից կատարում ենք գրադիենտային իջեցում` ստանալով y^2 կետը: գծ.3.14-ում տրված է այդ մեթոդի երկրաչափական մեկնաբանությունը:



Գծ. 3.14: Գելֆանդի մեթոդի երկրաչափական մեկնաբանությունը

Օրինակ: Գելֆանդի մեթոդով մինիմիզացնել $f(x) = 9x_1^2 + x_2^2$ ֆունկցիան: Որպես սկզնական մոտավորություն վերցնել $x^0 = (1,1)$ կետը: Կանգառի քայլ համարել $\|f'(x^k)\| \leq \varepsilon = 0.05$ պայմանը:

Վերցնենք $\tilde{x}^0 = (1.1, 1.1)$ և x^0, $\tilde{x}^0$ երկու կետերից կատարենք ամենաարագ գրադիենտային իջեցում` ստանալով $y^0 = (-0.008, 0, 888)$, $\tilde{y}^0 = (-0.009, 0.997)$ կետերը: Այնուհետև, կատարենք միաչափ մինիմիզացիա $\tilde{y^0} - y^0$ վեկտորի ուղղությամբ y^0 սկզբնակետով:

Ունենք

$$f(y^0 + \alpha(\tilde{y^0} - y^0)) = 9(-0.008 - 0.001\alpha)^2 +$$

$$+(0.888+0.089\alpha)^2=\varphi(\alpha):$$

Այսպեղից

$$argmin_{\alpha}\varphi(\alpha)=9.975,$$

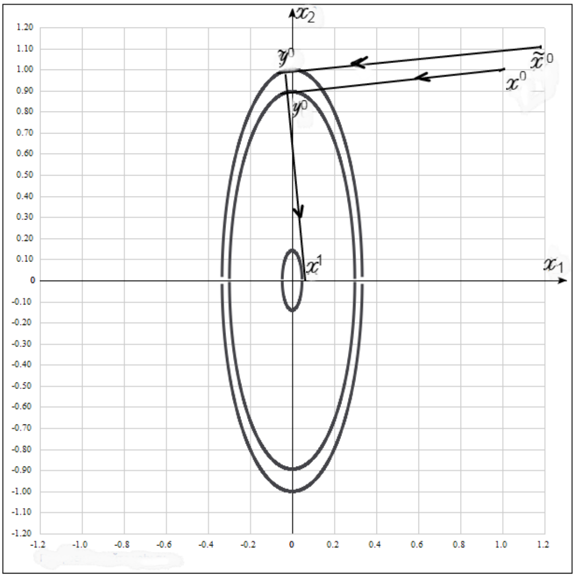
և հետևաբար

$$x^1=(0.002,0):$$

Քանի որ $\|f'(x^1)\|=0.036\le\varepsilon$, ապա

$$x_m=(0.002,0):$$

Այսպեղ նույնպես կապարվում է մեկ միաչափ մինիմիզացիայի քայլ պակաս, քան ամենաարագ վայրէջքի մեթոդում (տես գծ.3.15):



Գծ. 3.15: Գելֆանդի մեթոդը ($9x_1^2+x_2^2$ ֆունկցիայի դեպքը)

3.5 Ֆունկցիայի մինիմիզացիայի Նյուտոնի մեթոդը

Սովորաբար, երբ f ֆունկցիայի մակարդակի բազմություններ խիստ ձգված են ինչ որ մի ուղղությամբ, ապա նրա մինիմիզացիայի համար օգտագործում են, այսպես կոչված, երկրորդ կարգի մեթոդներ: Այս դեպքում, երբ $f''(x)$ հեսիանը դրական որոշյալ է և $f'(x)$ վեկտորը զրո չէ, ապա որպես ֆունկցիայի նվազման ուղղություն կարելի է վերցնել $-f''(x)^{-1}f'(x)$ վեկտորը: Գրականության մեջ սա հայտնի է որպես ֆունկցիայի մինիմիզացիայի Նյուտոնի մեթոդ և այն միաչափ դեպքի ընդհանրացումն է բազմաչափ դեպքի համար: Ենթադրվում է f ֆունկցիան երկու անգամ անընդհատ դիֆերենցելի է: Սկզբնական x^0 կետից սկսած կառուցում ենք անդրադարձ առնչություն հետևյալ բանաձևով.

$$x^{k+1} = x^k - (f''(x^k))^{-1}f'(x^k):$$

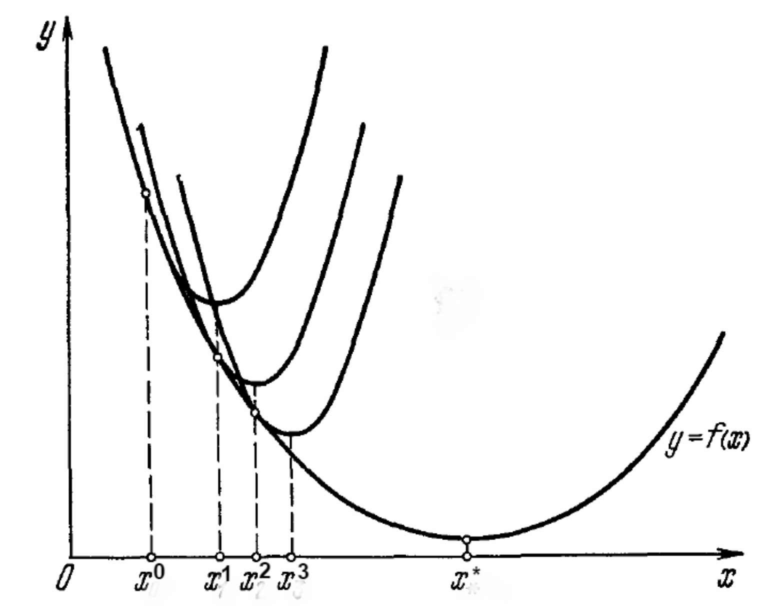
Տանք այս բանաձևի երկրաչափական իմաստը: Դիտարկենք քառակուսային հետևյալ ֆունկցիան.

$$\tilde{f}(x) \equiv f(x^k) + (f'(x^k), (x - x^k)) + \frac{1}{2}(f''(x^k)(x - x^k), x - x^k):$$

Այս քառակուսային ձևը ունի միակ մինիմում և այն

$$\tilde{f}'(x) = 0 = f'(x^k) + f''(x^k)(x - x^k)$$

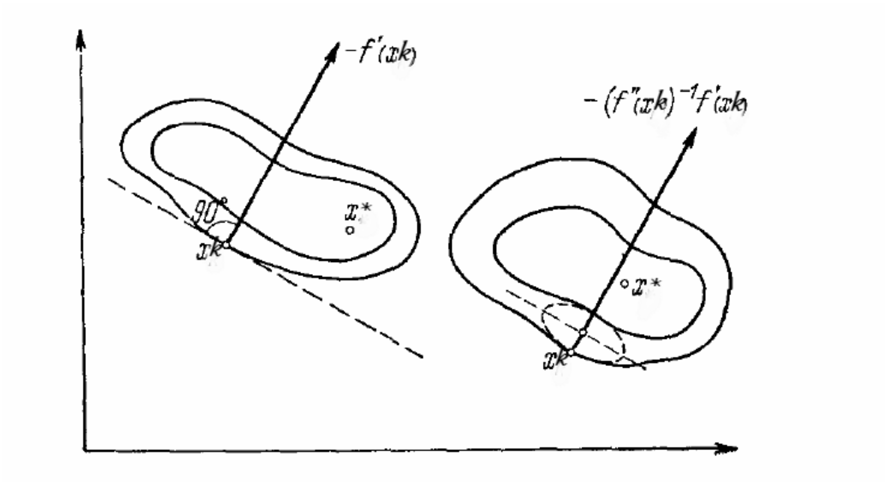
հավասարման լուծումն է, որը համընկնում է x^{k+1}-ի հետ: Հետևաբար Նյուտոնի մեթոդը մեկնաբանվում է որպես $\tilde{f}(x)$ փիպի քառակուսային ֆունկցիայի մինիմումի կետերի հաջորդական փնտրում: Երկրաչափորեն դա նշանակում է, որ x^k կետում f ֆունկցիան մոտարկվում է $\tilde{f}(x)$ պարաբոլով և որպես x^{k+1} հաջորդ մոտարկման կետ համարվում է այդ պարաբոլի մինիմումի կետը (տես գծ.3.16):



Գծ. 3.16:

Ինչպես տեսնում ենք այս մեթոդում քայլի երկարությունը յուրաքանչյուր իտերացիայում վերցվում է հաստատուն՝ հավասար մեկի: Այդ պատճառով մեթոդի զուգամիտության հարցը կախված է սկզբնական կետի ընտրությունից: Ուժեղ ուռուցիկ ֆունկցիաների համար նշվում է այն տիրույթը, որին եթե պատկանա սկզբնական կետը, ապա մեթոդը կլինի զուգամետ, իսկ զուգամիտության արագությունը կլինի քառակուսային (տես [24], Թեորեմ 2.1, գլուխ 5, պարագրաֆ 2, էջ 192): Գրադիենտային մեթոդի և Նյուտոնի մեթոդի համեմատությունը տրված է գծագիր 3.17-ում:

Ինչպես նկատելի է գրադիենտային իջեցման մեթոդում շարժումը կատարվում է f ֆունկցիայի մակարդակի բազմություններին ուղղահայաց, իսկ Նյուտոնի մեթոդում՝ մոտարկող քառակուսային $\tilde{f}$ ֆունկցիայի մակարդակի բազմություններին ուղղահաց: Այս դեպքում շարժումը ավելի է ուղղված մինիմումի կետի ուղղությամբ:



Գծ. 3.17: Գրադիենտային և Նյուտոնի մեթոդների համեմատությունը

Նյուտոնի մեթոդը մեկնաբանենք նախորդ պարագրաֆում դիտարկված օրինակի համար և համոզվենք, որ զուգամիտության պրոցեսը խիստ արագանում է: Դիցուք $f(x_1, x_2) = 9x_1^2 + x_2^2,\ x^0 = (1, 1)$:

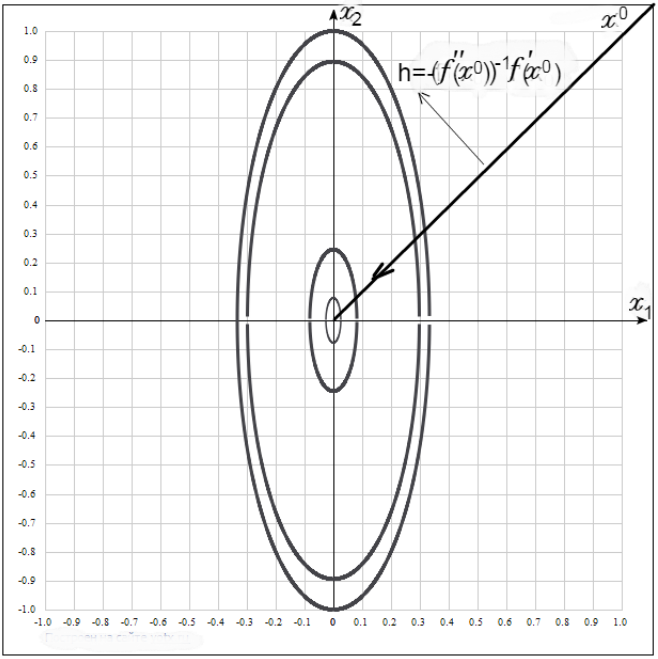
Ուստի

$$f'(x^0) = (18, 2), \quad f''(x^0)^{-1} = \begin{pmatrix} \frac{1}{18} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} :$$

Այսպեղից՝

$$h = -(f''(x^0))^{-1} f'(x^0) = (-1, -1):$$

Ուրեմն $x^1 = x^0 - h = (0,0)$: Հեփևաբար մինիմումի $(0,0)$ կեփը սփացվեց մեկ իփերացիայի ընթացքում (փես գծ.3.18): Հեշփ է համոզվել, որ ուրիշ սկզբնական կեփից սկսած այս մեթոդով նորից մեկ քայլի ընթացքում գալիս ենք $(0,0)$ մինիմումի կեփ: Այսինքն՝ այս օրինակի համար Նյուփոնի մեթոդը սկզբնական կեփի նկափմամբ կայուն է (զուգամեփ է բոլոր սկզբնական կեփերից սկսած):



Գծ. 3.18: $f(x_1, x_2) = 9x_1^2 + x_2^2$ ֆունկցիայի մինիմիզացիան երկ– րորդ կարգի մեթոդով

Նյուտոնի մեթոդի հիմնական թերությունը այն է, որ պետք է հաշվել ֆունկցիայի երկրորդ կարգի ածանցյալները, ինչը կարող է լինել էականորեն դժվար: Նշենք նաև, որ պրոցեսը կարող է փարամետ լինել, եթե նպատակային ֆունկցիան ուժեղ ուռուցիկ չէ կամ սկզբնական կետը բավականաչափ հեռու է մինիմումի կետից:

ԽՆԴԻՐՆԵՐ

Նյուտոնի երկրորդ կարգի մեթոդով լուծել հետևյալ խնդիրները և գրաֆիկորեն ցույց տալ իջեցման պրոցեսի շարժման հետագծերը:

1. $f(x) = x_1^2 + 3x_2^2 - x_1 x_2 + x_1 \to min$:

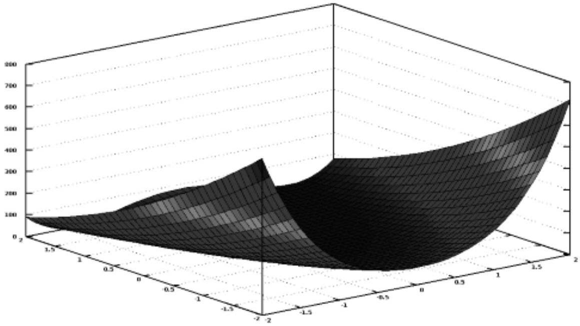
2. $f(x) = 2x_1^2 + x_2^2 + x_2 \to min$:

Առաջադրանք: Դիտարկենք հետևյալ ֆունկցիան.

$$f(x) = 100(x_2 - x_1^2)^2 + 5(1 - x_1)^2:$$

Այն գրականության մեջ հայտնի է Ռոզենբրոկի ֆունկցիա անվանմամբ: Նրա գրաֆիկը պատկերված է գծ. 3,19-ում:

Որպես սկզբնական մոտավորություն վերցնել $x^0 = (0,0)$ կետը: Օգտագործելով "$Matlab$" փաթեթը՝ մինիմիզացնել դիտարկվող ֆունկցիան ամենաարագ վայրէջքի, արագացված գրադիենտային, Գելֆանդի և քայլի կիսման մեթոդներով $\varepsilon = 0.003$ ճշտությամբ: Համեմատել այդ մեթոդները իտերացիաների քանակների տեսակետից: Նկարել յուրաքանչյուր մեթոդի շարժման հետագծերը:



Գծ. 3.19: Ռոզենբրոկի ֆունկցիայի գրաֆիկը

3.6 Համալուծ գրադիենտների մեթոդը

Դիտարկենք

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$$

քառակուսային ֆունկցիան, որտեղ $A(n \times n)$-ը սիմետրիկ դրական որոշյալ մատրից է, b-ն n չափանի վեկտոր է, իսկ c-ն թիվ է: Քանի որ $f(x)$-ը խիստ ուռուցիկ ֆունկցիա է, ապա արդեն գիտենք, որ այն ունի միակ մինիմումի կետ ամբողջ տարածության վրա: Նա հանդիսանում է $f'(x) = 0$ հավասարման լուծումը: Քանի որ $f'(x) = 0 \Leftrightarrow Ax + b = 0$, ապա $x_m = -A^{-1}b$: Այսինքն, եթե հայտնի է հակադարձ A^{-1} մատրիցը, ապա կարող ենք գտնել x_m մինիմումի կետը: Այժմ նկարագրենք մի մեթոդ, որն ամենաշատը n քայլերի ընթացքում գտնում է մինիմումի x^* կետը առանց մատրիցի հակադարձը հաշվելու: Դա համալուծ գրադիենտների մեթոդն է: Նկարագրենք այդ մեթոդը $n = 2$-ի դեպքում և տանք նրա երկրաչափական մեկնաբանությունը: f Ֆունկցիայի $V_C = \{x \in R^n / f(x) = C\}$ մակարդակի բազմությունները էլիպսներ են հարթության վրա: Վերցնում ենք կամայական սկզբնական x^0 մոտավորություն և կատարում ենք արագ վայրէջք $-f'(x^0)$ հակագրադիենտի ուղղությամբ՝ ստանալով $x^1 = x^0 - \alpha_0 f'(x^0)$ կետը: Մենք արդեն գիտենք, որ $f'(x^0)$ գրադիենտը ուղղահայաց է $V_{f(x_0)}$ մակարդակի բազմությանը: Նշանակենք $p^0 = -f'(x^0)$ և դիտարկենք

$$(Ap^0, x - x^1) = 0$$

գիծը հարթության վրա: Այն անցնում է x^1 կետով և նրա նորմալն է Ap^0 վեկտորը: Ցույց տանք, որ այդ գիծը անցնում է նաև x^* մինիմումի կետով: Իրոք, քանի որ $x^* = -A^{-1}b$, ապա

$$(Ap^0, x^* - x^1) = (p^0, A^T(x^* - x^1)) = (p^0, A(x^* - x^1)) =$$

$$= (p^0, -b - Ax^1) = (f'(x^0), f'(x^1)) = 0:$$

Այժմ, եթե այդ գծի վրա գտնենք մի p^1 վեկտոր և x^1 կետից p^1 վեկտորի ուղղությամբ կատարենք արագ վայրէջք, ապա արդյունքում կստանանք x^* մինիմումի կետը: p^1 վեկտորը կառուցենք հետևյալ բանաձևով.

$$p^1 = -\alpha f'(x^0) - f'(x^1):$$

α թիվն ընտրենք այնպես, որ այդ վեկտորը զուգահեռ լինի նշված գծին: Ուրեմն α-ն պետք է ընտրել

$$(Ap^0, -\alpha p^0 - f'(x^1)) = 0$$

հավասարությունից: Ուստի՝

$$\alpha = -\frac{(Ap^0, f'(x^1))}{(Ap^0, p^0)}:$$

Մեթոդի երկրաչափական մեկնաբանությունը տրված է գծ.3.20-ում: Այժմ նկարագրենք ընդհանուր ալգորիթմը:

- Սկզբնական x^0 կետում հաշվում ենք $-f'(x^0)$ և նշանակում՝ $p^0 = -f'(x^0)$:
- k-րդ քայլում որոշում ենք α_k-ն ամենաարագ վայրէջքի մեթոդով.

$$\alpha_k = argmin_{\alpha \geq 0} f(x^k + \alpha p^k):$$

  α_k-ի համար, համաձայն (3.1.2)-ի, ունենք հետևյալ բանաձևը.

$$\alpha_k = -\frac{(f'(x^k), p^k)}{(p^k, Ap^k)}:$$

  Կառուցում ենք x^k հաջորդականությունը ռեկուրենտ հետևյալ առնչությամբ.

$$x^{k+1} = x^k + \alpha_k p^k$$

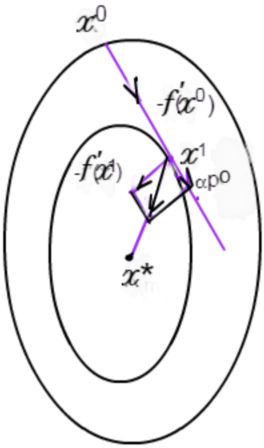
- Եթե $f'(x^{k+1}) = 0$, ապա ալգորիթմը ավարտվում է և x^{k+1}-ը մինիմումի կետն է: Հակառակ դեպքում p^{k+1} վեկտորը հաշվում ենք հետևյալ բանաձևով.

$$p^{k+1} = -f'(x^{k+1}) + \frac{(f'(x^{k+1}), f'(x^{k+1}))}{(f'(x^k), f'(x^k))} p^k$$

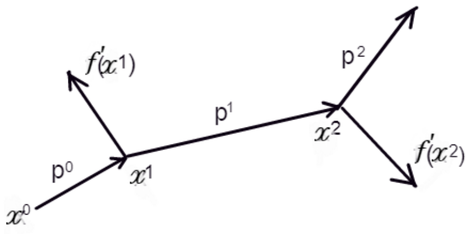
  և անցում կատարում հաջորդ իտերացիային:

$p^0, p^1, \ldots, p^n$ վեկտորները կոչվում են համալուծ գրադիենտներ: Այս ալգորիթմը գտնում է f-ի մինիմումի կետը ոչ ավելի քան n իտերացիաների ընթացքում: Գրականության մեջ այն հայտնի է որպես Ֆլետչեր-Ռիֆզի մեթոդ (տես [16,24]):

Համալուծ գրադիենտների մեթոդի երկրաչափական իմաստը հետևյալն է: Սկզբնական x_0 կետից կատարվում է արագ վայրէջք $p^0 = -f'(x^0)$ վեկտորի ուղղությամբ՝ սպանալով x^1 կետը: $f'(x^1)$ վեկտորը ուղղահայաց է p^0 վեկտորին: Այնուհետև ընտրվում է p^1 վեկտորը, որը համալուծ է p^0 վեկտորին՝ $(p^0, Ap^1) = 0$: Այնուհետև կատարվում է իջեցում p^1 վեկտորի ուղղությամբ և այսպես շարունակ (տես գծ.3.21):



Գծ. 3.20: Համալուծ գրադիենտների մեթոդի երկրաչափական մեկնաբանությունը



Գծ. 3.21: Համալուծ գրադիենտների երկրաչափական իմաստը

Օրինակ: Համալուծ գրադիենտների մեթոդով գտնել $f(x_1, x_2) = 2x_1^2 + x_2^2$ ֆունկցիայի մինիմումի կետը: Սկզբնական կետը համարել $x^0 = (2, 2)$ կետը: Կատարում ենք հետևյալ քայլերը:

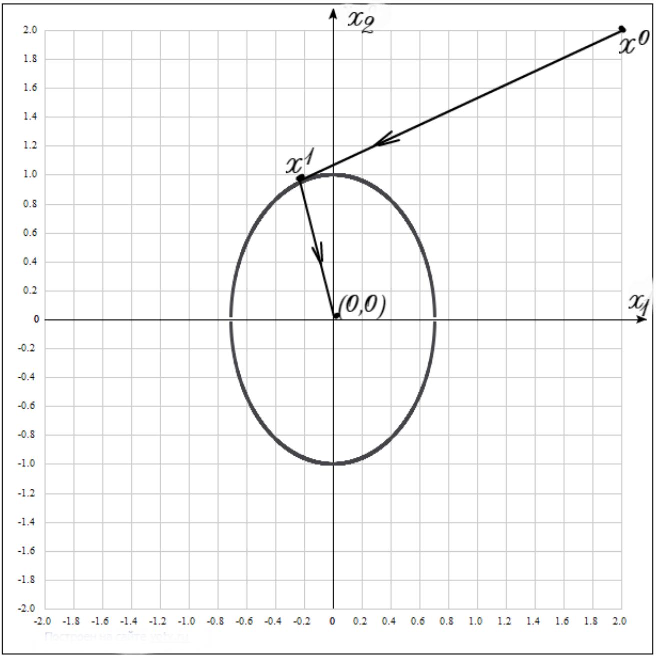
- Գտնում ենք ֆունկցիայի գրադիենտը x^0 կետում և կատարում ենք ամենաարագ վայրէջք $-f'(x^0)$ վեկտորի ուղղությամբ: Ունենք $f'(x^0) = (4x_1^0, 2x_2^0) = (8, 4)$: Ուստի $p^0 = -f'(x^0) = (-8, -4)$: Կատարելով ամենաարագ վայրէջք վեկտորի ուղղությամբ՝ կստանանք
$$x^1 = x^0 + \alpha_0 p^0,$$
որտեղ
$$\alpha_0 = \frac{(f'(x^0), f'(x^0))}{(Af'(x^0), f'(x^0))} = 0.278:$$
Ուրեմն $x^1 = (-0.22, 0.88)$:

- Այս քայլում կառուցում ենք p^0 վեկտորին համալուծ p^1 վեկտորը՝ օգտագործելով $(Ap^0, p^1) = 0$ պայմանը:

Այստեղ

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}:$$

Այստեղից $p^1 = (0.22, 0.88)$: Այնուհետև կատարելով արագ վայրէջք x^1 կետից p^1 վեկտորի՝ ուղղությամբ ստանում ենք $x^* = (0,0)$ մինիմումի կետը (տես գծ.3.22):



Գծ. 3.22: Համալուծ գրադիենտների մեթոդը $f(x_1, x_2) = 2x_1^2 + x_2^2$ ֆունկցիայի համար

ԽՆԴԻՐՆԵՐ

1. Համալուծ գրադիենտների մեթոդով գտնել $f(x_1, x_2) = x_1^2 + 2x_2^2 - 4x_1 - 4x_2$ ֆունկցիայի մինիմումի կետը:

2. Համալուծ գրադիենտների մեթոդով գտնել $f(x_1, x_2) = x_1^2 + 2x_2^2 + x_1x_2 + x_1 + x_2$ ֆունկցիայի մինիմումի կետը և գրաֆիկորեն ցույց տալ իջեցման պրոցեսի շարժման հետագիծը:

3.7 Քայլի կիսման եղանակի զուգամիտության թեորեմը

Լեմմ 3.7.1: *Դիցուք f ֆունկցիայի համար ճիշտ են հետևյալ պայմանները.*

1) f ֆունկցիայի գրադիենտը բավարարում է Լիպշիցի պայմանին, այսինքն՝ գոյություն ունի այնպիսի $L >$ > 0 հաստատուն, որ

$$\|f'(x) - f'(y)\| \leq L\|x - y\| \quad \forall x, y \in R^n:$$

2) f-ը ներքևից սահմանափակ է, այսինքն՝ գոյություն ունի այնպիսի $m > 0$ թիվ, որ տեղի ունի

$$f(x) \geq m \, \forall x \in R^n$$

անհավասարությունը:

Այդ դեպքում քայլի կիսման եղանակով վերջավոր քայլերի ընթացքում ընտրվում է α_k-ն և

$$\alpha_k > (1 - \varepsilon)/2L > 0, \quad k = 0, 1, 2, \dots \; :$$

▶ Դիցուք $h^k = -f'(x^k)$: Համաձայն Լագրանժի միջին արժեքի վերաբերյալ թեորեմի՝ գոյություն ունի այնպիսի $\theta \in (0,1)$ հաստատուն, որ եթե $\alpha > (1-\varepsilon)/L$, ապա

$$f(x^k + \alpha h^k) - f(x^k) = (f'(x^k + \alpha\theta h^k), \alpha h^k) =$$

$$= (f'(x^k + \alpha\theta h^k) - f'(x^k), \alpha h^k) + \alpha(f'(x^k), h^k) \leq$$

$$\leq \|L\alpha\theta h^k\|\|\alpha h^k\| - \alpha\|f'(x^k)\|^2 \leq L\alpha^2\|f'(x^k)\|^2 -$$

$$-\alpha\|f'(x^k)\|^2 = -\alpha\|f'(x^k)\|^2(1-\alpha L) \leq -\varepsilon\alpha\|f'(x^k)\|^2 \text{: } \blacksquare$$

Թեորեմ 3.7.1: *Դիցուք f ֆունկցիան բավարարում է* ***Լեմմ 3.7.1-ի*** *բոլոր պայմաններին և $\{x^k\}$-ն կիսման մեթոդով կառուցված հաջորդականությունն է:*

Այդ դեպքում $f'(x^k) \to 0$, երբ $k \to \infty$:

▶ Ունենք

$$f(x^{k+1}) - f(x^k) \leq -\varepsilon\alpha\|f'(x^k)\|^2 \text{:} \qquad (3.7.1)$$

Այստեղից հետևում է, որ $f(x^{k+1}) \leq f(x^k)$: Այսինքն՝ $\{f(x^k)\}$ հաջորդականությունը մոնոտոն նվազող է և ներքևից սահմանափակ է: Հետևաբար հաջորդականությունը զուգամետ է, այսինքն՝

$$f(x^k) - f(x^{k+1}) \to 0\text{:}$$

(3.7.1)-ից հետևում է, որ

$$\|f'(x^k)\|^2 \leq \frac{f(x^k) - f(x^{k+1})}{\varepsilon\alpha_k}\text{:}$$

Այստեղից, քանի որ $\alpha_k > (1-\varepsilon)/2L > 0$, ապա $f'(x^k) \to 0$: $\blacksquare$:

Դիտողություն: Նշենք, որ քայլի կիսման եղանակում էական դժվարությունը հիմնական (3.1.3) անհավասարության սպուգումն է: Եթե հայտի լինի L պարամետրը, ապա հարմար է օգտագործել հաստատուն քայլով գրադիենտային իջեցման եղանակը: Այստեղ բոլոր իտերացիաներում քայլի երկարությունը կարելի է վերցնել հաստատուն.

$$\alpha_k = \overline{\alpha} \equiv \frac{1-\varepsilon}{2L},\ k = 0, 1, \ldots:$$

Սակայն նշենք նաև, որ L հաստատունի արժեքը կախված է $f''(x)$ հեսիանի մաքսիմալ սեփական արժեքների վերևից հավասարաչափ ըստ x գնահատականի հետ, որն ընդհանուր դեպքում դժվար խնդիր է:

Թեորեմ 3.7.2: *Ենթադրենք գոյություն ունեն այնպիսի դրական $D > 0$, $d > 0$ հաստատուններ, որ*

$$D\|h\|^2 \geq (f''(x)h, h) \geq d\|h\|^2 \quad \forall x,\ h \in R^n : \qquad (3.7.2)$$

Այդ դեպքում կիսման եղանակով կառուցված $\{x^k\}$ հաջորդականությունը զուգամիտում է f-ի միակ x^* մինիմումի կետին երկրաչափական պրոգրեսիայի արագությամբ:

Այսինքն՝ գոյություն ունեն այնպիսի $C > 0$ և $q \in (0, 1)$ հաստատուններ, որ

$$\|x^k - x^*\| \leq Cq^k,\ k = 0, 1, ... :$$

► Քանի որ $f'(x^*) = 0$, ապա ըստ Թեյլորի բանաձևի

$$f(x) - f(x^*) = \frac{1}{2}(f''(x^* + \theta(x - x^*))(x - x^*), x - x^*),$$

որտեղ $\theta \in (0, 1)$: Այստեղից, հաշվի առնելով (3.7.2) անհավասարությունը, ստանում ենք

$$\frac{d}{2}\|x - x^*\|^2 \leq f(x) - f(x^*) \leq \frac{D}{2}\|x - x^*\| : \qquad (3.7.3)$$

Հաշվի առնելով այս պայմանը և ներկայացնելով f ֆունկցիան Թեյլորի բանաձևի տեսքով x կետում՝ կունենանք

$$f(x^*) - f(x) = (f'(x), x^* - x) +$$

$$+1/2(f''(x + \theta(x^* - x))(x^* - x), x^* - x) \geq$$
$$\geq -\|f'(x)\|\|x - x^*\| + d/2\|x - x^*\|^2:$$

Այստեղից կստանանք

$$f(x) - f(x^*) \leq \|f'(x)\|\|x - x^*\| - d/2\|x - x^*\|^2: \quad (3.7.4)$$

Հաշվի առնելով նաև (3.7.3) անհավասարության ձախ մասը՝ (3.7.4)-ից կստանանք

$$\frac{d}{2}\|x - x^*\|^2 \leq \|f'(x)\|\|x - x^*\| - d/2\|x - x^*\|^2:$$

Հետևաբար,

$$\|x - x^*\| \leq \frac{\|f'(x)\|}{d}: \quad (3.7.5)$$

Օգտվելով (3.7.3)-(3.7.5) անհավասարություններից՝ ստա– նում ենք

$$f(x) - f(x^*) \leq \frac{\|f'(x)\|^2}{d} - d/D(f(x) - f(x^*)):$$

Այստեղից

$$\|f'(x)\|^2 \geq d(1 + d/D)(f(x) - f(x^*)) : \quad (3.7.6)$$

Կիրառելով (3.7.6) անհավասարությունը կիսման մեթոդի

$$f(x^{k+1}) - f(x^k) \leq -\varepsilon\alpha_k\|f'(x^k)\|^2$$

անհավասարությունում՝ կստանանք

$$f(x^{k+1}) - f(x^*) \le [1 - \varepsilon\alpha_k d(1 + d/D)]((f(x^k) - f(x^*)):$$

Այստեղից, հաշվի առնելով $\alpha_k > \bar{\alpha} \equiv (1-\varepsilon)/L > 0$ անհավասարությունը, կստանանք

$$f(x^k) - f(x^*) \le (f(x^0) - f(x^*))\bar{q}^k, \qquad (3.7.7)$$

որտեղ

$$\bar{q} = 1 - \varepsilon\overline{\alpha}d(1 + d/D):$$

Վերջապես, օգտվելով (3.7.3) և (3.7.7) անհավասարությունից, կստանանք

$$\|x^k - x^*\| \le \sqrt{\frac{2}{d}}\sqrt{f(x^k) - f(x^*)} \le Cq^k,$$

որտեղ

$$C = \sqrt{\frac{2}{d}}\sqrt{f(x^0) - f(x^*)},\ q = \sqrt{\bar{q}} : \blacksquare$$

Քայլի կիսման եղանակի պրակտիկ իրականացումների ժամանակ սովորաբար օգտվում են նրա հետևյալ մոդիֆիկացված տարբերակից:
k-րդ քայլում որպես ֆունկցիայի նվազման ուղղություն վերցնում են

$$h^k = -\frac{f'(x^k)}{\|f'(x^k)\|}$$

վեկտորը, իսկ α_k քայլի երկարությունը ընտրվում է կիսման եղանակի հետևյալ պայմանից.

$$f(x^k + \alpha h^k) - f(x^k) \le -\alpha\varepsilon\|f'(x^k)\|: \qquad (3.7.8)$$

Հաջորդ x^{k+1} կետը կառուցվում է

$$x^{k+1} = x^k + \alpha_k h^k \tag{3.7.9}$$

ռեկուրենտ առնչությամբ:
Այժմ քայլի կիսման եղանակը մեկնաբանենք հետևյալ պարզ օրինակի միջոցով:

Օրինակ: Քայլի կիսման եղանակով գտնել

$$f(x_1, x_2) = 4x_1^2 + 4x_2^2 + 6x_1x_2$$

ֆունկցիայի մինիմումի կետը R^2-ի վրա: Որպես սկզբնական մոտավորություն վերցնել $x^0 = (-2, 1)$ կետը: ε պարամետրի արժեքը վերցնել հավասար 0.5-ի: Կատարել իտերացիայի մեկ քայլ:

Լուծում: Քայլի կիսման եղանակի իտերացիայի (3.7.9) բանաձևը այս խնդրի համար ունի հետևյալ տեսքը.

$$x^{k+1} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \end{pmatrix} = \begin{pmatrix} x_1^k \\ x_2^k \end{pmatrix} + \alpha_k \begin{pmatrix} h_1^k \\ h_2^k \end{pmatrix} : \tag{3.7.10}$$

Մասնակի ածանցյալների և գրադիենտի նորմի համար ունենք հետևյալ բանաձևերը.

$$f'_{x_1}(x_1, x_2) = 8x_1 + 6x_2, \quad f'_{x_2}(x_1, x_2) = 8x_2 + 6x_1, \tag{3.7.11}$$

$$\|f'(x)\| = \sqrt{(8x_1 + 6x_2)^2 + (8x_2 + 6x_1)^2}, \tag{3.7.12}$$

$$h_1^k = -\frac{8x_1^k + 6x_2^k}{\|f'(x^k)\|}, \ h_2^k = -\frac{6x_1^k + 8x_2^k}{\|f'(x^k)\|} : \tag{3.7.13}$$

Առաջին իտերացիայում $k = 0$:

Օգտագործելով (3.7.10) – (3.7.13) բանաձևերը՝ կստանանք

$$f'_{x_1}(x_1^0, x_2^0) = 8x_1^0 + 6x_2^0 = 8(-2) + 6 = -10,$$

$$f'_{x_2}(x_1^0, x_2^0) = 6x_1^0 + 8x_2^0 = -4,$$

$$\|f'(x^0)\| = \sqrt{(-10)^2 + (-4)^2} \approx 10.77, \; h_1^0 = \frac{10}{10.77} \approx 0.93,$$

$$h_2^0 = \frac{4}{10.77} \approx 0.37:$$

Դիցուք $\alpha = 1$: Այդ դեպքում, հաշվի առնելով այս արդյունքները, մեկնարկային $(-2, 1)$ արժեքը և (3.7.10)-ը, կստանանք

$$x \equiv \begin{pmatrix} -2 \\ 1 \end{pmatrix} - \begin{pmatrix} 0.93 \\ 0.37 \end{pmatrix} = \begin{pmatrix} -1.07 \\ 1.63 \end{pmatrix}:$$

Առաջին իտերացիայի համար α_0 քայլի ընտրության պայմանը ունի հետևյալ տեսքը.

$$f(x) - f(x^0) \leq -0.5\alpha\|f'(x^0)\|:$$

Կատարելով համապատասխան հաշվարկներ՝ նկատում ենք, որ այս անհավասարության ձախ մասը մոտավորապես հավասար է -1.6-ի, իսկ աջ մասը հավասար է -5.4-ի: Հետևաբար այն տեղի չունի: Այժմ կիսենք α թիվը և նորից ստուգենք անհավասարությունը: Այս դեպքում

$$x = (-1.58, 1.18):$$

Քանի որ $f(x^0) = 8$, $f(x) \approx 4.16$, ապա անհավասարության ձախ մասը հավասար է $4.16 - 8 = -3.84$, իսկ աջ մասը, հեշտ է նկատել, որ հավասար է -2.69-ի:

Այսպիսով, նշված անհավասարությունը փեղի ունի և հե–պևաբար՝
$$\alpha_0 = 0.5,\ x^1 = (-1.54, 1.18):$$

ԽՆԴԻՐՆԵՐ

Քայլի կիսման եղանակով լուծել հետևյալ էքստրեմումի խնդիրները: Վերցնել $x^0 = (1, 1),\ \varepsilon = 0.5$:
Կապարել իտերացիայի մեկ քայլ:

1. $2x_1^2 + 2x_2^2 + 4x_1 + 6x_2 - 2x_1x_2 \to min$:

2. $-2x_1^2 - 2x_2^2 - 2x_1x_2 + 2 \to max$:

3.8 Գծայնացման մեթոդը

Դիտարկենք պայմանական օպտիմիզացիայի հետևյալ խնդիրը.
$$f(x) \to min,\ x \in M: \tag{3.8.1}$$
Այս խնդրում f-ը դիֆերենցելի ուռուցիկ ֆունկցիա է, իսկ $M \subset R^n$-ը կոմպակտ ուռուցիկ բազմություն է:

Տանք գծայնացման մեթոդի համառոտ նկարագրությունը: Ընտրում ենք կամայական x^0 կետ M բազմությունից և կառուցում ենք $\{x^k\}$ հաջորդականությունը
$$x^{k+1} = x^k + \alpha_k h^k,\ k = 0,\ 1, ...$$
ռեկուրենտ առնչությամբ: Այստեղ h^k վեկտորը f-ի այնպիսի նվազման ուղղություն է $x^k \in M$ կետում, որ բավականաչափ փոքր α_k քայլերի դեպքում $x^{k+1} \in M$: h^k վեկտորի որոշման համար k-րդ քայլում լուծում ենք հետևյալ միջանկյալ խնդիրը.
$$(f'(x^k), x - x^k) \to min,\ x \in M:$$

Ենթադրենք $\bar{x}^k$-ն այդ խնդրի որևէ լուծում է: Նշանակենք

$$\eta_k = (f'(x^k), \bar{x}^k - x^k),\ h^k = \bar{x}^k - x^k:$$

h^k վեկտորը կոչվում է պայմանական հակագրադիենտ: Ակնհայտ է, որ $\eta_k \leq 0$: Իրոք

$$\eta_k = \min_{x \in M}(f'(x^k), x - x^k) \leq (f'(x^k), x^k - x^k) \leq 0:$$

Եթե $\eta_k < 0$, ապա ընտրում ենք α_k քայլը կիսման մեթոդով: Վերցնում ենք $\alpha = 1$ և ստուգում

$$f(x^k + \alpha h^k) - f(x^k) \leq \alpha \varepsilon \eta_k \tag{3.8.2}$$

անհավասարությունը: Եթե այն տեղի ունի, ապա համարում ենք $\alpha_k = 1$, հակառակ դեպքում α-ն կիսում ենք և նորից ստուգում նշված անհավասարությունը և այսպես շարունակ: Երբ առաջին անգամ տեղի ունենա (3.8.2) անհավասարությունը, ապա այդ α-ն համարվում է α_k-ի արժեք և ցիկլը ավարտվում է: Այնուհետև կառուցում ենք x^{k+1} կետը ռեկուրենտ առնչությամբ.

$$x^{k+1} = x^k + \alpha_k h^k:$$

Թեորեմ 3.8.1: *Դիցուք*

ա) $M \subset R^n$-ը ուռուցիկ կոմպակտ է,

բ) $f(x)$-ը դիֆերենցելի է և ուռուցիկ M-ի վրա,

գ) $f'(x)$ գրադիենտը M բազմության վրա բավարարում է Լիպշիցի պայմանին:

Այդ դեպքում

1) *եթե որևէ* k*-երորդ քայլում* $\eta_k = 0$*, ապա* x^k*-ն (3.8.1) խնդրի լուծումն է և ալգորիթմն ավարտվում է,*

2) *եթե* $\eta_k < 0,\ k = 0, 1, 2, ...,$ *ապա*

$$f(x^k) \to \min_{x \in M} f(x):$$

► Դիցուք $\eta_k = 0$: Ըստ ուռուցիկ ֆունկցիայի հիմնական անհավասարության՝ ունենք

$$f(x) - f(x^k) \geq (f'(x^k), x - x^k) \geq \eta_k = 0 \ \forall x \in M,$$

այսինքն՝ x^k-ն (3.8.1) խնդրի լուծումն է:

Ցույց տանք, որ h^k ուղղությամբ բավականաչափ փոքր շարժվելիս մնում ենք M բազմության մեջ: Իրոք, քանի որ M-ը ուռուցիկ է, ապա

$$x^k + \alpha h^k = x^k + \alpha(\bar{x}^k - x^k) = (1-\alpha)x^k + \alpha\bar{x}_k \in M, \ \forall \alpha \in [0, 1]:$$

Քանի որ f-ի գրադիենտը M կոմպակտի վրա բավարարում է Լիպշիցի պայմանին և $x_k \in M$, ապա կարելի է ցույց տալ, որ վերջավոր քայլերից հետո (3.8.2) անհավասարությունը տեղի ունի և հետևաբար α_k քայլը ընտրվում է: Միաժամանակ գոյություն ունի այնպիսի $\bar{\alpha} > 0$ թիվ, որ $\alpha_k > \bar{\alpha} > 0,\ k = 0, 1, ...$ (տես, օրինակ՝ [14], Լեմմ 3.1, էջ 229):
Այժմ ապացուցենք, որ

$$f(x^k) \to \min_{x \in M} f(x):$$

Համաձայն (3.8.2) անհավասարության՝ ունենք

$$f(x^{k+1}) \leq f(x^k) + \varepsilon\alpha_k\eta_k: \tag{3.8.3}$$

Գումարելով (3.8.3) անհավասարությունները $k \in [0 : m-1]$ ինդեքսների համար՝ կստանանք

$$\min_{x\in M} f(x) - f(x^0) \leq f(x^m) - f(x^0) \leq \sum_{k=0}^{m-1} \alpha_k \eta_k:$$

Այստեղից հետևում է, որ բացասական անդամներով $\sum \alpha_k \eta_k$ շարքը զուգամետ է: Հետևաբար, նրա ընդհանուր անդամը ձգտում է զրոյի՝

$$\alpha_k \eta_k \to 0: \tag{3.8.4}$$

Քանի որ $\alpha_k > \bar{\alpha} > 0$, ապա (3.8.4)-ից հետևում է, որ $\eta_k \to 0$: Այստեղից, ընտրելով $x^{k_j} \to x^* \in M$ զուգամետ ենթահաջորդականությունը և օգտվելով ուռուցիկ ֆունկցիայի հիմնական անհավասարությունից, կստանանք՝

$$f(x) - f(x^{k_j}) \geq (f'(x^{k_j}), x - x^{k_j}) \geq \eta_{k_j} \ \forall x \in M: \tag{3.8.5}$$

Քանի որ $\eta_{k_j} \to 0$ և $x^{k_j} \to x^*$, ապա (3.8.5) անհավասարությունում անցնելով սահմանի՝ կստանանք՝

$$f(x) - f(x^*) \geq 0 \ \ \forall x \in M,$$

այսինքն՝ x^*-ը f-ի մինիմումի կետն է M բազմության վրա: Այսպիսով, ապացուցվեց, որ $\{f(x^k)\}$ հաջորդականության $\{f(x^{k_j})\}$ ենթահաջորդականությունը զուգամիտում է $\min_{x\in M} f(x)$: Մյուս կողմից, քանի որ $\{f(x^k)\}$ հաջորդականությունը մոնոտոն նվազող է, ապա այն նույնպես կզուգամիտի $\min_{x\in M} f(x)$-ին: ■

Շատ դեպքերում α_k քայլի ընտրության համար կատարվում է $f(x)$ ֆունկցիայի մինիմիզացիա $[x^k, \overline{x}^k]$

հատվածի վրա, ինչպես հետևյալ երկու օրինակներում: **Օրինակ:** Գծայնացման մեթոդով լուծել հետևյալ խնդիրը.

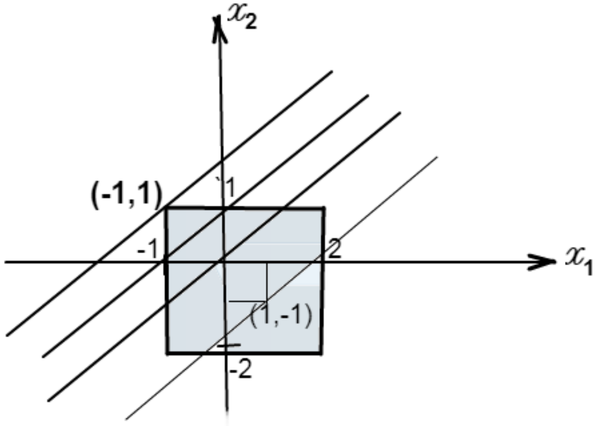
$$\begin{cases} x_1^2 + x_2^2 \to min, \\ -1 \leq x_1 \leq 2, \\ -2 \leq x_2 \leq 1 : \end{cases},$$

Սկզբնական մոտավորությունն է $x^0 = (1, -1)$ կետը: Ունենք $f'(x^0) = (2, -2)$:

Կազմենք միջանկյալ գծային խնդիրը.

$$\begin{cases} 2(x_1 - 1) - 2(x_2 + 1) \to min, \\ -1 \leq x_1 \leq 2, \\ -2 \leq x_2 \leq 1 : \end{cases}$$

Այս խնդրի լուծումը $\overline{x} = (-1, 1)$ կետն է (տես գծ.3.23):



Գծ. 3.23: Միջանկյալ գծային խնդրի լուծումը

Գտնենք $\alpha_0 = \min\{1, \alpha_*\}$ թիվը, որտեղ

$$\alpha_* = argmin_{\alpha > 0} f(x^0 + \alpha(\overline{x} - x^0)):$$

Քանի որ

$$x^0 + \alpha(\overline{x} - x^0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 - 2\alpha \\ -1 + 2\alpha \end{pmatrix}:$$

Այսպեղից կսպանանք

$$\alpha_* = argmin((1 - 2\alpha)^2 + (-1 + 2\alpha)^2) = \frac{1}{2} < 1:$$

Հեպևաբար

$$\alpha_0 = \alpha_*, \ x^1 = x^0 + \alpha_0(\overline{x} - x^0) = (0, 0):$$

Ուրեմն, $\eta_1 = 0$, ինչը նշանակում է, որ x^1 կեպը խնդրի լուծումն է:

Օրինակ: Գծայնացման մեթոդով գտնել $f(x_1, x_2) = (x_1 - 2)^2 + x_2^2$ ֆունկցիայի մինիմումի կեպը

$$M = \{(x_1, x_2) \in R^2 / -1 \leq x_1 \leq 1, \ -1 \leq x_2 \leq 1\}$$

բազմության վրա: Որպես սկզբնական մոտավորություն վերցնել $x^0 = (-1, 1)$ կեպը:

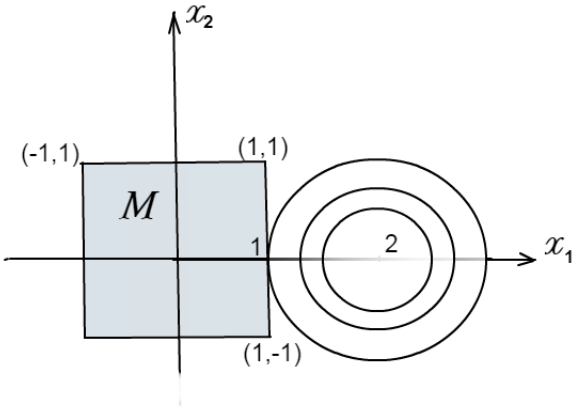
Քանի որ f ֆունկցիայի մակարդակի բազմությունները համակենտրոն շրջանագծեր են ընդհանուր $(2, 0)$ կենտրոնով, ապա պարզ է, որ $(1, 0)$ կեպը խնդրի լուծումն է (տես գծ.3.24):

Այժմ քայլ առ քայլ նկարագրենք գծայնացման մեթոդը՝ տալով երկրաչափական մեկնաբանություններ:

- Առաջին քայլում հաշվում ենք $f'(x^0)$ գրադիենտը և մինիմիզացնում ենք $(f(x^0), x - x^0)$ գծային ֆունկցիան M բազմության վրա: Ունենք $f'(x^0) = (-4, 2)$: Հեպևաբար

$$(f(x^0), x - x^0) = -4(x_1 + 1) + 2(x_2 - 1)$$

:



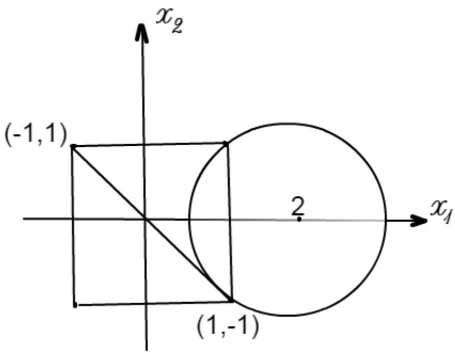
Գծ. 3.24: խնդրի լուծումը գրաֆիկորեն

- Գտնում ենք $-4(x_1+1)+2(x_2-1)$ գծային ֆունկցիայի մինիմումի կետը M բազմության վրա: Երկրաչափորեն լուծելով այս խնդիրը՝ ստանում ենք $\overline{x}^0 = (1, -1)$ կետը:

- Գտնում ենք f ֆունկցիայի մինիմումի կետը $[x^0, \overline{x}^0]$ հատվածի վրա: Լուծելով այս խնդիրը երկրաչափորեն՝ ստանում ենք $x^1 = \overline{x}^0 = (1, -1)$ (տես գծ.3.25):

- Այնուհետև մինիմիզացնելով $(f'(x^1), x - x^1)$ ֆունկցիան M բազմության վրա՝ ստանում ենք $\overline{x}^1 = (1, 1)$ կետը:

- Այս քայլում գտնում ենք f ֆունկցիայի մինիմումի կետը $[x^1, \overline{x}^1]$ հատվածի վրա՝ ստանում ենք $x^2 = (1, 0)$ կետը, որը ընդհանուր խնդրի լուծումն է:



Գծ. 3.25: Հաջորդ կետի կառուցումը գրաֆիկորեն

ԽՆԴԻՐՆԵՐ

1. Ապացուցել պնդումները հետևյալ խնդրի համար.

$$(c, x) \to min,\ x \in M:$$

ա) Եթե $M = \{x \in R^n / \|x - x^0\| \leq r\}$,
ապա

$$x^* = x^0 + \frac{c}{\|c\|} r$$

վեկտորը խնդրի լուծումն է:

բ) Եթե $M = \{x \in R^n / a_j \leq x_j \leq b_j,\ j \in [1:n]\}$,
ապա

$$x_j^* = \begin{cases} a_j, & \text{եթե } c_j \geq 0, \\ b_j, & \text{եթե } c_j < 0 \end{cases}$$

կոորդինատներով վեկտորը խնդրի լուծումն է:

2. Գտնել $f(x) = x_1^2 + x_2^2$ ֆունկցիայի պայմանական հակագրադիենտը $x = (2,3)$ կետում $M = \{x \in \in R^2 / x_1 + x_2 \le 6,\ x_1 \ge 0,\ x_2 \ge 0\}$ բազմության վրա:

3. Գծայնացման մեթոդով լուծել հետևյալ խնդիրը.
$$\begin{cases} x_1^2 + x_2^2 \to min, \\ (x_1 - 2)^2 + (x_2 - 2)^2 \le 8 : \end{cases}$$
Որպես սկզբնական մոտավորություն վերցնել $x^0 = (2,2)$ կետը:

Առաջադրանք: Կազմել ծրագիր, որը իրականացնում է $f(x) = 1/2(\mathbf{A}x, x) + (b, x)$ ֆունկցիայի մինիմիզացիան $M = \{x \in R^n / \mathbf{C}x \le d,\ x \ge 0\}$ բազմության վրա պայմանական գրադիենտի մեթոդով: Այստեղ $\mathbf{A}$-ն $(n \times n)$ չափանի սիմետրիկ դրական որոշյալ մատրից է, իսկ $\mathbf{C}$-ն $(m \times n)$ չափանի մատրից է: k-րդ քայլում օգտագործելով սիմպլեքս ալգորիթմը՝ ստանալ

$$\eta_k = \min_{x \in M} (f'(x^k), x - x^k)$$

խնդրի որևէ լուծում: Ալգորիթմի կանգառի համար ընդունել $|\eta_k| < \varepsilon_0$ պայմանը, որտեղ $\varepsilon_0 > 0$ նախապես տրված ճշտություն է: Եթե նշված պայմանը կատարվում է, ապա x^k վեկտորը համարել խնդրի լուծում և ավարտել ալգորիթմը: Սկզբնական $x^0 \in M$ կետի ընտրությունը նույնպես կատարել սիմպլեքս ալգորիթմով:

3.9 Ապրիորի մեթոդի զուգամիտությունը

Թեորեմ 3.9.1: *Դիցուք f-ը ուռուցիկ ֆունկցիա է՝ որոշված R^n-ի վրա և M^*-ը նրա մինիմումի կետերի* *բազմություններ է R^n-ի վրա: Ենթադրենք $M^* \neq \emptyset$: Դիցուք $\{x^k\}$-ն հետևյալ ռեկուրենտ առընչությամբ կառուցված հաջորդականություն է.*

$$x^0 \in R^n,\ x^{k+1} = x^k - \alpha_k \frac{f'(x^k)}{\|f'(x^k)\|},\quad k = 0, 1, 2, ..., \quad (3.9.1)$$

որտեղ α_k-երը դրական թվեր են՝ բավարարող

$$\sum \alpha_k = +\infty,\ \sum \alpha_k^2 < +\infty \qquad (3.9.2)$$

պայմաններին:

Այդ դեպքում

$$x^k \to \bar{x} \in M^*:$$

Այսինքն՝ $\{x^k\}$ հաջորդականությունը զուգամիտում է f-ի որևէ $\overline{x}$ մինիմումի կետի:

▶ Նշանակենք $v^k = f'(x^k)$: Վերցնենք որևէ $x^* \in M^*$ կետ և հաշվենք նրա հեռավորությունը $\{x^k\}$ հաջորդականության անդամներից:
Ունենք

$$\|x^{k+1} - x^*\|^2 = \|x^k - x^*\|^2 + \frac{2\alpha_k}{\|v^k\|}(v^k, x^* - x^k) + \alpha_k^2: \quad (3.9.3)$$

Քանի որ x^*-ը f-ի մինիմումի կետ է M-ի վրա, ապա

$$(v^k, x^* - x^k) \le f(x^*) - f(x^k) \le 0:$$

Հաշվի առնելով այս պայմանը և (3.9.3)-ը՝ կստանանք

$$\|x^{k+1} - x^*\|^2 \le \|x^k - x^*\|^2 + \sum_{i=0}^{m-1} \alpha_i^2: \qquad (3.9.4)$$

Քանի որ $\sum {\alpha_k}^2$ շարքը զուգամետ է, ապա (3.9.4) անհավասարությունից հետևում է, որ $\{x^k\}$-ն սահմանափակ հաջորդականություն է: Այստեղից հետևում է, որ սահմանափակ կլինի նաև $\{v^k\}$ հաջորդականությունը: Այսինքն՝ գոյություն ունի այնպիսի $C > 0$ թիվ, որ

$$\|v^k\| \leq C: \qquad (3.9.5)$$

Ցույց տանք, որ գոյություն ունի ինդեքսների այնպիսի $\{k_s\}$ ենթահաջորդականություն, որ

$$(v^{k_s}, x^* - x^{k_s}) \to 0, \text{ երբ } k_s \to \infty: \qquad (3.9.6)$$

Ենթադրենք հակառակը: Այդ դեպքում գոյություն ունի այնպիսի $N > 0$ թիվ, որ ինչ-որ K համարից սկսած տեղի ունի

$$(f'(x^k), x^* - x^k) < -N < 0, \; \forall k > K \qquad (3.9.7)$$

անհավասարությունը:

Օգտվելով (3.9.3)-(3.9.7) անհավասարություններից՝ կըստանանք

$$\|x^{k+1} - x^*\|^2 \leq \|x^0 - x^*\|^2 - \frac{2N}{C}\sum_{i=0}^{k}\alpha_i + \sum_{i=0}^{k}\alpha_i^2: \quad (3.9.8)$$

Այս անհավասարության մեջ անցնելով սահմանի, երբ $k \to \infty$, կստանանք, որ նրա աջ մասը ձգտում է $-\infty$, իսկ ձախ մասը ոչ բացասական թիվ է, ինչը հակասություն է: Այժմ $\{x_{k_s}\}$ ենթահաջորդականության կետերի համար, կիրառելով ուռուցիկ ֆունկցիայի հիմնական անհավասարությունը, կունենանք

$$0 \geq f(x^*) - f(x^{k_s}) \geq (v^{k_s}, x^* - x^{k_s}):$$

Այսպեղ անցնելով սահմանի, երբ $k_s \to \infty$ և հաշվի առնելով նաև (3.9.6)-ը, կստանանք

$$f(x^{k_s}) \to f(x^*) = \min_{x \in M} f(x):$$

Քանի որ $\{x^{k_s}\}$-ը սահմանափակ հաջորդականություն է, ապա նրանից կարելի է անջատել զուգամետ ենթահաջորդականություն: Ընդհանրությունը չխախտելով, ենթադրենք, որ $x^{k_s} \to \bar{x}$: Այսպեղից, օգտվելով f-ի անընդհատությունից, կստանանք

$$f(x^{k_s}) \to f(\bar{x}) = f(x^*) = \min_{x \in M} f(x),$$

այսինքն՝ $\bar{x} \in M^*$: Ցույց տանք, որ $x^k \to \bar{x}$: Քանի որ $x^{k_s} \to \bar{x}$, ապա կամայական $\varepsilon > 0$ թվի համար գոյություն ունի այնպիսի K համար, որ երբ $k_s > K$, ապա տեղի ունի

$$\|x^{k_s} - \bar{x}\|^2 < \frac{\varepsilon}{2} \tag{3.9.9}$$

անհավասարությունը: Կարող ենք նաև համարել, որ ցանկացած p համարի համար տեղի ունի

$$\sum_{i=k_s}^{k_s+p-1} \alpha_i^2 < \frac{\varepsilon}{2} \tag{3.9.10}$$

անհավասարությունը: Օգտվելով (3.9.3) և (3.9.9)-(3.9.10) անհավասարություններից՝ ստանում ենք, որ երբ $k_s > K$, ապա ցանկացած p բնական թվի համար տեղի ունի հետևյալ անհավասարությունը.

$$\|p^{k_s+p} - \bar{x}\|^2 \le \|x^{k_s} - \bar{x}\|^2 + \sum_{i=k_s}^{k_s+p-1} \alpha_i^2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon:$$

Իսկ սա նշանակում է, որ

$$x^k \to \bar{x}: \blacksquare$$

ԽՆԴԻՐՆԵՐ

Ապրիորի մեթոդով լուծել հետևյալ խնդիրները:
Վերցնել $x^0 = (1, 1),\ \alpha_k = 1/k + 1, k = 0, 1, 2, ...$:
Կատարել իտերացիայի երկու քայլ:

1. $3x_1^2 + x_2^2 + 11x_2 + 3x_1 \to min$:

2. $-2x_1^2 - x_2^2 + 8x_1 + 6x_2 - 25 \to max$:

3.10 Գրադիենտի պրոյեկտման մեթոդը

Դիցուք $M \subseteq R^n$-ը փակ ուռուցիկ բազմություն է: Նշանակենք $\Pi_M(a)$-ով a վեկտորի **պրոյեկցիան** M բազմության վրա: Այսինքն՝ $\Pi_M(a)$-ն M բազմության ամենամոտիկ կետն է a վեկտորից:

Ճշմարիտ է հետևյալ պնդումը (տես, օրինակ՝ [24], Լեմմ 2.1, էջ 225):

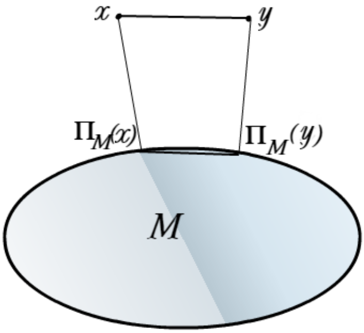
Լեմմ 3.10.1: Π_M *օպերատորը բավարարում է Լիպշիցի պայմանին* $L = 1$ *հաստատունով, այսինքն՝*

$$\|\Pi_M(x) - \Pi_M(y)\| \le \|x - y\| \ \forall\, x,\ y \in R^n \qquad (3.10.1)$$

(տես գծ.3.26):

Դիտարկենք մաթեմատիկական ծրագրավորման հետև–յալ խնդիրը.

$$f(x) \to min,\ x \in M:$$



Գծ. 3.26: Պրոյեկցման օպերատորը ուռուցիկ կոմպակտի վրա (սեղմելու հատկությունը)

Այստեղ $f(x)$-ը դիֆերենցելի ֆունկցիա է, իսկ M-ը ուռուցիկ բազմություն է: Գրադիենտի պրոյեկման մեթոդով այս խնդրի լուծման համար կառուցվում է $\{x_k\}$ հաջորդականություն հետևյալ ռեկուրենտ առնչությամբ.

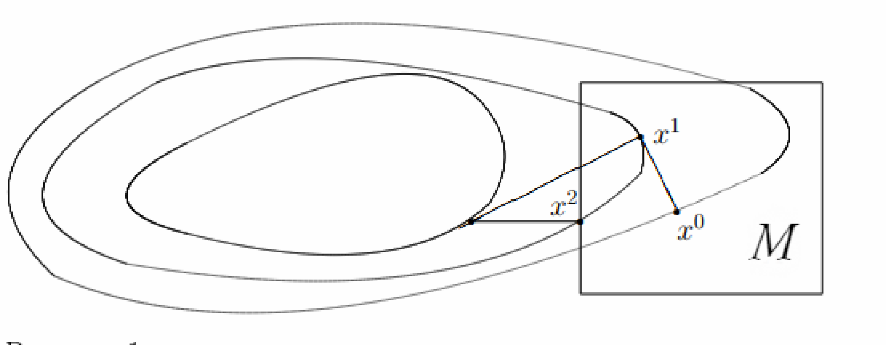
$$x^0 \in M, \ \ x^{k+1} = \Pi_M(x^k - \alpha_k f'(x^k)), \ k = 0, 1, 2, ...$$

(տես գծ.3.27):

Ուսումնասիրենք այն պայմանները, որոնց դեպքում այս հաջորդականությունը կզուգամիտի f-ի որևէ մինիմումի կետի:

Լեմմ 3.10.2: *Դիցուք f-ը θ հաստատունով ուժեղ ուռուցիկ ֆունկցիա է M ուռուցիկ բազմության վրա: Այդ դեպքում տեղի ունի հետևյալ անհավասարությունը.*

$$(f'(x) - f'(y), x - y) \geq 2\theta\|x - y\|^2 \quad \forall\, x,\ y \in M: \quad (3.10.2)$$



Գծ. 3.27:

► Ըստ f ֆունկցիայի ուժեղ ուռուցիկության (1.1.2) սահմանման կամայական $x,\ y \in M$ կետերի համար ունենք

$$f(x) - f(y) \geq (f'(x), x - y) + \theta\|x - y\|^2,$$

$$f(y) - f(x) \geq (f'(y), y - x) + \theta\|x - y\|^2:$$

Գումարելով այս երկու անհավասարությունները՝ կստանանք (3.10.2) անհավասարությունը: ■

Հենվելով այս արդյունքների վրա՝ ապացուցենք պրոյեկման մեթոդի զուգամիտության հետևյալ թեորեմը:

Թեորեմ 3.10.1: *Դիցուք f-ը θ հաստատունով ուժեղ ուռուցիկ ֆունկցիա է M փակ ուռուցիկ բազմության վրա: Ենթադրենք նաև, որ f' գրադիենտը բավարարում է Լիպշիցի պայմանին M բազմության վրա $L > 0$ հաստատունով, այսինքն՝*

$$\|f'(x) - f'(y)\| \leq L\|x - y\| \; \forall x,\ y \in M:$$

Այդ դեպքում, եթե պրոյեկտման մեթոդում որպես α_k քայլեր վերցնենք միևնույն հաստատունը $(0, \frac{4\theta}{L^2})$

միջակայքից, ապա $\{x^k\}$ հաջորդականությունը կզուգամիտի f-ի միակ մինիմումի կետին երկրաչափական պրոգրեսիայի արագությամբ:

▶ Դիտարկենք հետևյալ օպերատորը՝ $\mathbf{A}_\alpha : M \to M$,

$$\mathbf{A}_\alpha(x) \equiv \Pi_M(x - \alpha f'(x)):$$

Ցույց տանք, որ այն սեղմող օպերատոր է: Օգտագործելով (3.10.1) – (3.10.2) անհավասարությունները՝ կունենանք.

$$\|\mathbf{A}_\alpha(x) - \mathbf{A}_\alpha(y)\|^2 = \|\Pi_M(x - \alpha f'(x)) - \Pi_M(y - \alpha f'(y))\|^2 \leq$$

$$\leq \|x - \alpha f'(x) - y + \alpha f'(y)\|^2 \leq \|x - y + \alpha(f'(y) - f'(x))\|^2 =$$
$$= \|x - y\|^2 + 2\alpha(x - y, f'(y) - f'(x)) + \alpha^2\|f'(x) - f'(y)\|^2 =$$
$$\leq \|x - y\|^2(1 - 4\theta\alpha + \alpha^2 L^2): \qquad (3.10.3)$$

Ընտրենք α այնպես, որ

$$q \equiv \sqrt{1 - 4\alpha\theta + \alpha^2 L^2}$$

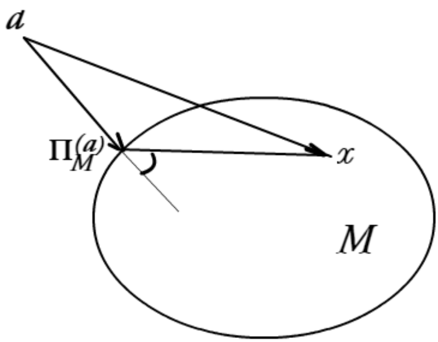
թիվը փոքր լինի մեկից: Եթե $\alpha \in (0, \frac{4\theta}{L^2})$, ապա $q < 1$ և հետևաբար $\mathbf{A}_\alpha$ օպերատորը կլինի սեղմող: Դիցուք x^*-ը նրա անշարժ կետն է M բազմության վրա, այսինքն՝

$$\Pi_M(x^* - \alpha f'(x^*)) = x^* : \qquad (3.10.4)$$

Ցույց տանք, որ այդ անշարժ կետը f-ի մինիմումի կետն է: Իրոք, քանի որ պրոյեկտման օպերատորը բավարարում է

$$(\Pi_M(a) - a, x - \Pi_M(a)) \geq 0 \;\forall a \in R^n,\; \forall x \in M \qquad (3.10.5)$$

անհավասարությանը (տես, օրինակ՝ [24], լեմմ 2.1, էջ 225, տես. նաև գծ. 3.28: α անկյունը սուր է), ապա այստեղ



Գծ. 3.28: Կետի պրոյեկցիան ուռուցիկ բազմության վրա ու նրա հապկությունը

տեղադրելով $a = x^* - \alpha f'(x^*)$, $\Pi_M(a) = x^*$ և հաշվի առնելով (3.10.4)-ը, կստանանք՝

$$(x^* - x^* + \alpha f'(x^*), x - x^*) \geq 0 \Rightarrow$$

$$\Rightarrow (f'(x^*), x - x^*) \geq 0 \;\; \forall\, x \in M : \qquad (3.10.6)$$

Վերջապես, հաշվի առնելով նաև ուռուցիկ ֆունկցիայի հիմնական անհավասարությունը, (3.10.6)-ից կունենանք

$$f(x) - f(x^*) \geq (f'(x^*), x - x^*) \geq 0 \;\forall x \in M,$$

այսինքն՝ x^*-ը f-ի մինիմումի կետն է:

Բացի դրանից, (2.5.3)-ից հետևում է, որ

$$\|x^{k+1} - x^*\| = \|\mathbf{A}_\alpha x^k - \mathbf{A}_\alpha x^*\| \leq q\|x^k - x^*\| \leq ...$$

$$\leq q^{k+1}\|x^0 - x^*\|:$$

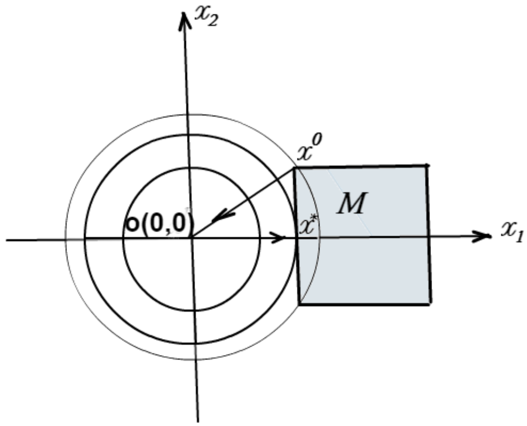
Իսկ սա նշանակում է, որ $\{x^k\}$ հաջորդականությունը զուգամիտում է x^* կետին երկրաչափական պրոգրեսիայի արագությամբ: ■

Օրինակ: Դիցուք պետք է պրոյեկտման մեթոդով գտնել $f(x_1,x_2)=x_1^2+x_2^2$ ֆունկցիայի մինիմումի կետը

$$M=\{(x_1,x_2)\in R^2/1\le x_1\le 3,\ -1\le x_2\le 1\}$$

բազմության վրա: Պարզ է, որ $(1,0)$-ը մինիմումի կետն է: Եթե որպես սկզբնական մոտավորություն վերցնենք $x^0=(1,1)$ կետը, ապա այդ կետից ամենաարագ վայրէջքով գալիս ենք շրջանի կենտրոն՝ $(0,0)$ կետը, որը M բազմությունից դուրս է: Այնուհետև այն պրոյեկտելով M բազմության վրա՝ ստանում ենք խնդրի լուծումը՝ $x^*=(1,0)$ կետը:

Գծագիր 3.29-ում տրված է այս պրոցեսի երկրաչափական մեկնաբանությունը:



Գծ. 3.29:

ԽՆԴԻՐՆԵՐ

1. Դիցուք $a \in R^n$ կամայական կետ է: Ապացուցել, որ եթե

ա) $M = \{x \in R^n / \|x - x^0\| \leq r\}$,
ապա

$$\Pi_M(a) = x^0 + \frac{a - x^0}{\|a - x^0\|} r,$$

բ) եթե $M = \{x \in R^n / b_j \leq x_j \leq c_j,\ j \in [1:n]\}$,
ապա

$$(\Pi_M(a))_j = \begin{cases} b_j, & \text{եթե } a_j < b_j \\ a_j, & \text{եթե } b_j \leq a_j \leq c_j \\ c_j, & \text{եթե } a_j > c_j, \end{cases}$$

գ) եթե $M = \{x \in R^n / x_j \geq 0,\ j \in [1:n]\}$,
ապա

$$\Pi_M(a) = (max(0, a_1),\ ...,\ max(0, a_n)):$$

2. Գրադիենտի պրոյեկտման մեթոդով գտնել $f(x_1, x_2) = 2x_1^2 + x_2^2 + x_1 x_2$ ֆունկցիայի մինիմումի կետը

$$M = \{(x_1, x_2) \in R^2 / x_1 + 2x_2 \leq -8\}$$

բազմության վրա: Որպես սկզբնական մոտավորություն վերցնել $x^0 = (-5, -3)$ կետը:

3. Գրադիենտի պրոյեկտման մեթոդով գտնել $f(x_1, x_2) = x_1^2 + x_2^2$ ֆունկցիայի մինիմումի կետը

$$M = \{(x_1, x_2) \in R^2 / -7 \leq x_1 \leq 5,\ -5 \leq x_2 \leq -2\}$$

բազմության վրա: Որպես սկզբնական մոտավորություն վերցնել $x^0 = (-6, -1)$ կետը:

Առաջադրանք: Կազմել ծրագիր, որը գրադիենտի պրո–յեկտման մեթոդով կիրականացնի

$$f(x) = 1/2(\mathbf{A}x, x) + (b, x)$$

քառակուսային ֆունկցիայի մինիմիզացիան գնդի և զուգահեռանիստի վրա: Որպես պրոյեկտման օպերատորներ վերցնել վերևում նշված բանաձևերը: L հաստատունը վերցնել $\mathbf{A}$ մատրիցի նորմը, ուժեղ ուռուցիկության θ հաստատունը վերցնել մատրիցի մինիմալ սեփական արժեքը, իսկ

$$\alpha_k = \frac{2\theta}{L^2}:$$

Կանգառի քայլ համարել $\|x^{k+1} - x^k\| < \varepsilon_0$ պայմանը, որտեղ $\varepsilon_0 > 0$ նախապես տրված ճշտություն է:

3.11 Տուգանքային ֆունկցիաների մեթոդը

Դիտարկենք պայմանական օպտիմիզացիայի ընդհանուր խնդիրը.

$$f(x) \to min, x \in M:$$

Սահմանենք հետևյալ ֆունկցիան՝

- $H(x) > 0 \ \forall x \in R^n$,
- $H(x) = 0 \Leftrightarrow x \in M$,
- H-ը անընդհատ ֆունկցիա է:

Այս ֆունկցիան կոչվում է տուգանքային ֆունկցիա: Դիտարկենք r պարամետրից կախված $\varphi(x,r) = f(x) + rH(x)$ ֆունկցիան R^n-ի վրա: Տեղի ունի հետևյալ թեորեմը ([24], Թեորեմ 5.1, էջ 242):

Թեորեմ 3.11.1: *Դիցուք*

- $M \subset R^n$*-ը կոմպակտ բազմություն է և*

$$H(x) \to \infty, \text{ երբ } \|x\| \to \infty :$$

- $f(x)$*-ը անընդհատ է և* $f(x) \to \infty$*, երբ* $\|x\| \to \infty$*:*
- $r_{k+1} > r_k, \ r_k \to \infty$*:*

Այդ դեպքում $x(r_k) = argmin_{x \in R^n}(f(x) + r_k H(x))$ *հաջորդականության սահմանային կետը խնդրի լուծումն է և* $H(x(r_k)) \to 0$*:*

Եթե X բազմությունը տրվում է

$$X = \{x \in R^n / g_i(x) \le 0, \ i = 1, 2, \dots, m,$$

$$h_j(x) = 0, j = 1, 2, \dots, l\}$$

տեսքով, որտեղ $g_i, \ i = 1, 2 \dots, m, \ h_j, \ j = 1, 2, \dots, l$ ֆունկցիաները անընդհատ դիֆերենցելի են, ապա, սովորաբար, որպես տուգանք վերցնում են հետևյալ ֆունկցիան.

$$H(x) = \sum_{i=1}^{m} [max(0, g_i(x))]^2 + \sum_{j=1}^{l} h_j^2(x):$$

Այս ֆունկցիան անընդհատ դիֆերենցելի է: Իրոք, առաջին գումարի յուրաքանչյուր անդամի համար կարող ենք գրել՝

$$[max(0, g_i(x))]^2 = \begin{cases} g_i^2(x), \text{ եթե } g_i(x) \geq 0, \\ 0, \text{ եթե } g_i(x) < 0 \end{cases} :$$

Այստեղից

$$([max(0, g_i(x))]^2)' = \begin{cases} 2g_i'(x)g_i(x), \text{ եթե } g_i(x) \geq 0, \\ 0, \text{ եթե } g_i(x) < 0 \end{cases} =$$

$$= 2g_i'(x)max(0, g_i(x)) :$$

Հետևաբար

$$H'(x) = \sum_{i=1}^{m} 2g_i'(x)max(0, g_i(x)) + \sum_{j=1}^{l} 2h_j'(x)h_j(x) :$$

Այսպիսով, միջանկյալ ոչ պայմանական օպտիմիզացիայի խնդրի լուծման համար կարելի է օգտագործել գրադիենտային իջեցման մեթոդներից որևէ մեկը:

Այժմ նկարագրենք տուգանքային ֆունկցիաների մեթոդի ընդհանուր ալգորիթմը:

- Վերցնում ենք մոնոտոն աճող դրական թվերի այնպիսի $\{r_k\}$ հաջորդականություն, որ

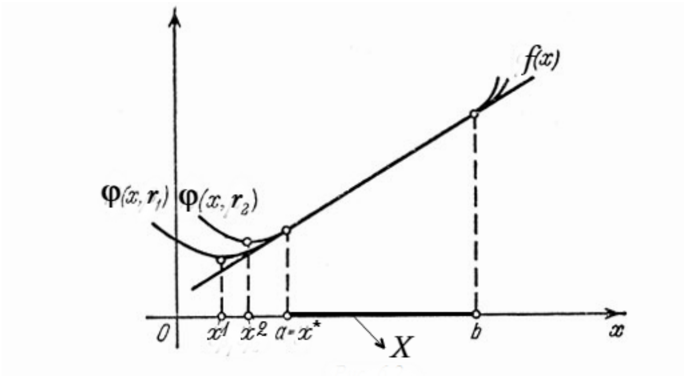
  $$r_{k+1} > r_k, \ r_k \to \infty:$$

  Դիցուք $x(r_0) = x^0$ և $\varepsilon > 0$:

- $x^0 = x(r_{k-1}), \ x(r_k) = argmin_{x \in R^n}(f(x) + r_k H(x))$:

- Եթե $H(x(r_k)) < \varepsilon$, ապա կապարում ենք կանգառ և ընդունում՝ $x_m = x(r_k)$: Հակառակ դեպքում k-ն ավելացնում ենք մեկով և վերադառնում նախորդ քայլին:

Գծ.3.30-ում պրված է այս մեթոդի երկրաչափական մեկ– նաբանությունը:



Գծ. 3.30: Տուգանքային ֆունկցիաների մեթոդի երկրաչափական մեկնաբանությունը

Օրինակ: Տուգանքային ֆունկցիաների մեթոդով լուծել հետևյալ խնդիրը.

$$\begin{cases} f(x) = x_1^2 + x_2^2 \rightarrow min, \\ g(x) = 2x_1 + x_2 + 4 \leq 0 : \end{cases}$$

Վերցնենք $x^0 = (0,0), r_1 = 1, \ r_{k+1} = 10r_k, \ \varepsilon = 0.01$: Միջանկյալ ոչ պայմանական օպտիմիզացիայի խնդիրները լուծենք $\varepsilon_1 = 0.001$ ճշտությամբ: Ունենք

$$\varphi(x, r_1) = f(x) + H(x) =$$

$$= x_1^2 + x_2^2 + [max(0, 2x_1 + x_2 + 4)]^2, \ x^{10} = x^0:$$

Նշանակենք x^{ij}-ով այն կետը, որը ստացվում է j իտերացիայում ոչ պայմանական օպտիմիզացիայի i խնդրի լուծման արդյունքում ամենաարագ վայրէջքի մեթոդով նշված ճշտությամբ: Այսպիսով,

$$\varphi'(x, r_1) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} + 2max(0, 2x_1 + x_2 + 4) \begin{pmatrix} 2 \\ 1 \end{pmatrix}:$$

Այստեղից

$$x^{11} = x^{10} - \alpha_0 \varphi'(x^0, r_1) = \begin{pmatrix} 1 - 30\alpha_0 \\ 1 - 16\alpha_0 \end{pmatrix},$$

որտեղ $\alpha_0 = argmin \varphi(\alpha, r_1)$,

$$\varphi(\alpha, r_1) = (1 - 30\alpha)^2 + (1 - 16\alpha)^2 +$$

$$+[max(0, 7 - 76\alpha)]^2:$$

Հետևաբար

$$\alpha_0 = 0.083, \ x^{11} = (-1.501, -0.334),$$

$$\|\varphi'(x^{11}, 1)\| > \varepsilon_1 :$$

Շարունակելով անալոգ ձևով՝ կստանանք

$$x^{12} = \begin{pmatrix} -1.326 \\ -0.661 \end{pmatrix}, \ \|\varphi'(x^{12}, 1)\| > \varepsilon_1,$$

$$x^{13} = \begin{pmatrix} -1.333 \\ -0.665 \end{pmatrix}, \ \|\varphi'(x^{13}, 1)\| > \varepsilon_1,$$

$$x^{14} = \begin{pmatrix} -1.333 \\ -0.666 \end{pmatrix}, \ \|\varphi'(x^{14}, 1)\| < \varepsilon_1:$$

Այսպիսով, $x(r_1) = x^{14}$: Կանգառի պայմանը դեռևս տեղի չունի, քանի որ

$$H(x(r_1)) = 0.444 > \varepsilon:$$

Ուստի անցնենք երկրորդ իտերացիային: Ընդունենք $x^{20} = x(r_1)$ և կառուցենք

$$\varphi(\alpha, r_2) = f(x) + 10H(x) =$$

$$= x_1^2 + x_2^2 + 10[max(0, 2x_1 + x_2 + 4)]^2$$

ֆունկցիան: Երկրորդ իտերացիայում կստանանք

$$x^{21} = \begin{pmatrix} -1.568 \\ -0.784 \end{pmatrix}, \ \|\varphi'(x^{21}, 10)\| < \varepsilon_1$$

Հետևաբար $x(r_2) = x^{21}$,

$$H(x(r_2)) = 0.006 < \varepsilon:$$

Ուրեմն կանգառի պայմանը կատարվում է: Ուստի ընդունում ենք, որ $x_m = x(r_2)$ և ալգորիթմը ավարտում ենք:

ԽՆԴԻՐՆԵՐ

Տուգանքային ֆունկցիաների մեթոդով լուծել հետևյալ խնդիրները.

1.

$$\begin{cases} f(x) = x^2 \to min \\ g(x) = x - 1 \le 0, \end{cases},$$

$$x_0 = 5, r_1 = 1, \ r_{k+1} = 10r_k, \ \varepsilon = 0.01:$$

2.

$$\begin{cases} f(x) = x_1 + x_2 \to min, \\ x_1^2 + x_2^2 \le 1, \end{cases},$$

$$x_0 = 0.5, r_1 = 1,\ r_{k+1} = 10r_k,\ \varepsilon = 0.1:$$

Վերջում նշենք, որ պայմանական օպտիմիզացիայի թվային մեթոդները և ալգորիթմները բուռն զարգացում ստացան միայն այն ժամանակ, երբ ստեղծվեցին էլեկտրոնային հաշվիչ մեքենաները: Դա կապված էր այդ խնդիրների հաշվողական դժվարությունների հետ: 20-րդ դարի 60-ական թվականներից սկսած այդ բնագավառի վերաբերյալ բուռն ուսումնասիրություններ կատարվեցին աշխարհի տարբեր մաթեմատիկական դպրոցներում: Օրինակ, տուգանքային ֆունկցիաների վերաբերյալ լայնածավալ և ամբողջական ուսումնասիրություններ կատարեցին ամերիկացի մաթեմատիկոսներ Է. Ֆիակկոն և Գ. Մակ-Կորմիկը: Նրանց մենագրությունը ոչ գծային ծրագրավորման թվային մեթոդների վերաբերյալ համարվում է դասական (տես [28]): Գրադիենտի պրոյեկտման մեթոդը առաջին անգամ առաջարկվել է ամերիկացի մաթեմատիկոս Ռոզենի կողմից 1960 թվականին: Գծայնացման մեթոդի վերաբերյալ սիստեմատիկ և ամբողջական ուսումնասիրություն կատարվել է Ուկրաինայի ազգային ակադեմիայի կիբեռնետիկայի ինստիտուտում Բ. Ն. Պշենիչնիի կողմից 1970-1980 ական թվականներին: Բավական է նշել միայն նրա մենագրությունը գծայնացման մեթոդի վերաբերյալ(տե՛ս [23]):

Գլուխ 4

Ուռուցիկ անալիզ

Ուռուցիկ անալիզը մաթեմատիկական անալիզի բաժին է, որն ուսումնասիրում է ուռուցիկ բազմությունների և ուռուցիկ ֆունկցիաների հատկությունները:

Ուռուցիկ անալիզի գաղափարները և փաստերը ֆունդամենտալ նշանակություն ունեն օպտիմիզացիայի թվային մեթոդների տեսությունում: Դրանք լայն կիրառություններ ունեն նաև կիրառական մաթեմատիկայի այնպիսի բնագավառներում ինչպիսիք են՝ խաղերի տեսությունը, գործույթների հետազոտումը, մաթեմատիկական էկոնոմիկան և այլն:

Այս գլուխը կարելի է դիտարկել որպես ուռուցիկ անալիզի ներածություն: Այստեղ բերված փաստերը և պնդումները օգտագործվում են հետագայում մաթեմատիկական ծրագրավորման խնդիրներում Լագրանժի անորոշ գործակիցների մեթոդը հիմնավորելու համար: Նշենք, որ կան նաև լրացուցիչ փաստեր, որոնք ձևակերպված են խնդիրների տեսքով:

Հարկ ենք համարում նշել, որ ուռուցիկ անալիզի տեսության սկիզբը դրվել է Մինկովսկու(1891) և Ֆենխելի(1949) աշխատանքներում, իսկ ուռուցիկ անալիզի ժամանակակից մեթոդների հետ ընթերցողը կարող է ծանոթանալ [18, 22, 26, 30] աշխատանքներում:

4.1 Ուռուցիկ բազմությունների անջատման թեորեմները

Սահմանում 4.1.1: *Դիցուք $p \in R^n$-ն ոչ զրոյական վեկտոր է, իսկ α-ն իրական թիվ է:*

$$H = \{x \in R^n / (p, x) = \alpha\}$$

բազմությունը կոչվում է հիպերհարթություն, իսկ

$$H_+ = \{x \in R^n / (p, x) \geq \alpha\}, \ \ H_- = \{x \in R^n / (p, x) \leq \alpha\}$$

բազմությունները՝ այդ հիպերհարթությամբ ծնված կիսատարածություններ:

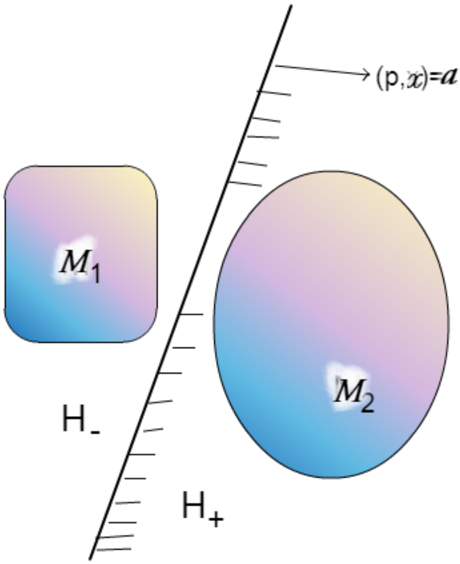
Սահմանում 4.1.2: *$M_1, \ \ M_2 \subseteq R^n$ բազմությունները կոչվում են **անջատվող հիպերհարթությամբ**, եթե գոյություն ունի այնպիսի $p \neq 0$ վեկտոր, որ*

$$(p, x) \leq (p, y) \ \forall x \in M_1, \ \forall y \in M_2:$$

Սա նշանակում է, որ գոյություն ունի այնպիսի H հիպերհարթություն, որ

$$M_1 \subseteq H_-, \ \ M_2 \subseteq H_+:$$

Գծ.4.1-ում պրված է երկու ուռուցիկ բազմությունների անջատման երկրաչափական մեկնաբանությունը հարթության վրա:



Գծ. 4.1: Ուռուցիկ բազմությունների անջատումը հիպերհար–թությամբ

Թեորեմ 4.1.1: *Դիցուք $M_1,\ M_2 \subseteq R^n$ այնպիսի ուռուցիկ բազմություններ են, որ $M_1 \cap M_2 = \emptyset$: Այդ դեպքում նրանք անջատվում են հիպերհարթությամբ:*

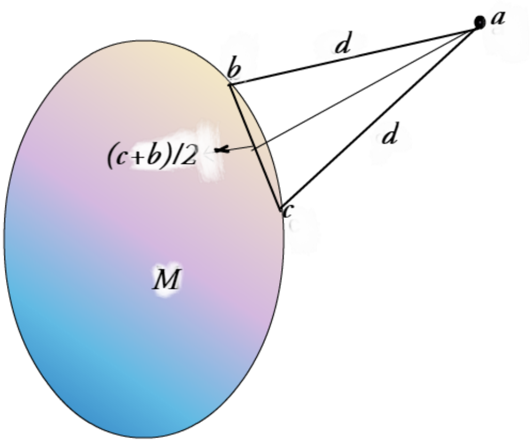
► Թեորեմի ապացույցը հենվում է հետևյալ երկու պնդումների վրա:

Լեմմ 4.1.1: *Եթե M-ը փակ ուռուցիկ բազմություն է և $a \notin M$, ապա այդ կետը հիպերհարթությամբ անջատվում է M բազմությունից:*

► Գտնենք a կետի պրոյեկցիան M բազմության վրա:

Ցույց փանք, որ այն գոյություն ունի և միակն է: Վերցնենք a կենտրոնով և $r \equiv inf_{x\in M}\|a-x\|+\varepsilon$ շառավղով $B_r(a)$ գունդը, որտեղ $\varepsilon > 0$ ֆիքսած թիվ է: Քանի, որ գունդը փակ սահմանափակ բազմություն է, ապա $M \cap B_r(a)$ հատումը կոմպակտ է: Հետևաբար, a կետի հեռավորությունը այդ հատումից հասանելի է: Դժվար չէ համոզվել, որ այդ կետը ամենամոտիկն է նաև M բազմությունից, այսինքն՝ դա $\Pi_M(a)$ կետն է: Ցույց տանք $\Pi_M(a)$-ի միակությունը: Ենթադրենք գոյություն ունի երկու ամենամոտիկ կետ: Դիցուք դրանք b, c կետերն են (տես գծ.4.2): Եթե a, b, c կետերը եռանկյուն են կազմում, ապա այն հավասարասրուն է, որի սրունքը հավասար է $d \equiv \inf_{x\in M}\|x-a\|$, իսկ հիմքը $[c, b]$ հատվածն է: Քանի որ $M-$ ը ուռուցիկ է, ապա $[c, b] \in M$: Հետևաբար a գագաթից տարած բարձրության հիմքը $[c, b]$ հատվածի միջնակետն է: Քանի որ բարձրությունը փոքր է սրունքից և սրունքը մինիմալ հեռավորությունն է, ապա ստացանք հակասություն: Նշենք նաև a, b, c կետերը միևնույն գծի վրա գտնվել չեն կարող, քանի որ $a \notin M$:

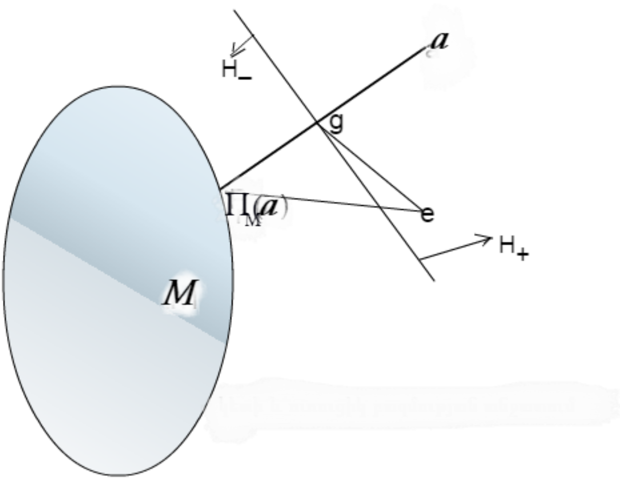
Այժմ $[\Pi_M(a), a]$ հատվածի g միջնակետով տանենք հիպերհարթություն, որի նորմալը $a - \Pi_M(a)$ վեկտորն է (տե՛ս գծ.4.3): Ցույց տանք, որ այդ հիպերհարթությունը անջատում է a կետը M բազմությունից: Դիցուք H_+-ը այն կիսատարածությունն է, որը պարունակում է a կետը, իսկ H_- կիսատարածությունը չի պարունակում a-ն: Ցույց տանք, որ $M \subset H_-$: Նախ պարզ է, որ g կետի պրոյեկցիան M բազմության վրա $\Pi_M(a)$ կետն է: Ենթադրենք, որ գոյություն ունի $e \in M \cap H_+$: Այդ դեպքում g, $\Pi_M(a), e$ գագաթներով եռանկյան մեջ g գագաթից տարված բարձրության հիմքը կպատկանա $[\Pi_M(a), e] \subseteq M$ հատվածին և այդ բարձրությունը փոքր է $[g, \Pi_M(a)]$ հատվածի երկարությունից, որը հակասություն է, քանի որ



Գծ. 4.2: Պրոյեկցիայի միակությունը

այն g կետի մինիմալ հեռավորությունն է M բազմությունից:
■

Լեմմ 4.1.2: *Դիցուք M−ը ուռուցիկ բազմություն է, իսկ $a \notin M$: Այդ դեպքում a կետը հիպերհարթությամբ անջատվում է M բազմությունից:*
► Եթե $a \notin \overline{M}$, ապա հանգում ենք նախորդ դեպքին, քանի որ այս դեպքում a կետը կարելի է հիպերհարթությամբ անջատել $\overline{M}$ բազմությունից, հետևաբար՝ M բազմությունից: Այժմ ենթադրենք, որ a-ն M-ի եզրային կետ է: Այդ դեպքում գոյություն ունի այնպիսի $\{x^k\} \notin \overline{M}$ հաջորդականություն, որ $x^k \to a$: Համաձայն լեմմ 4.1.1-ի յուրաքանչյուր x^k կետ կարելի է անջատել M բազմությունից հիպերհարթությամբ: Դա նշանակում է, որ գոյություն ունեն այնպիսի p^k միավոր



Գծ. 4.3: Կետի և ուռուցիկ բազմության անջատումը հիպեր–հարթությամբ

վեկտորներ, որ

$$(p^k, x) \le (p^k, x^k) \ \ \forall x \in M:$$

Չսահմանափակելով ընդհանրությունը, կարելի է ենթադրել, որ $p^k \to p^0 \neq 0$: Անցնելով սահմանի՝ կստանանք

$$(p^0, x) \le (p^0, a) \ \forall x \in M,$$

ինչը նշանակում է a կետի և M բազմության անջատում հիպերհարթությամբ: ■
Այժմ անցնենք թեորեմի ապացույցին: Նշանակենք $M \equiv M_1 - M_2$: M-ը ուռուցիկ է և քանի որ $M_1 \cap M_2 = \emptyset$,

ապա $0 \notin M$: Համաձայն Լեմմ 4.1.2-ի՝ 0 կետը կարելի է անջատել M բազմությունից: Դա նշանակում է, որ գոյություն ունի այնպիսի $p \neq 0$ վեկտոր, որ

$$(p, x - y) \leq 0 \; \forall x \in M_1, \; \forall y \in M_2 : \blacksquare$$

Այստեղից

$$(p, x) \leq (p, y) \; \forall x \in M_1, \forall y \in M_2:$$

Հետևանք 4.1.1: *Դիցուք $M \subseteq R^n$-ը ուռուցիկ բազմություն է, իսկ a-ն նրա եզրային կետը: Այդ դեպքում a կետով կարելի է տանել այնպիսի հիպերհարթություն, որ M բազմությունը ընկած լինի այդ հիպերհարթությամբ ծնված կիսատարածություններից որևէ մեկում: Այդպիսի հիպերհարթությունը կոչվում է **հենման** հիպերհարթություն:*

Սահմանում 4.1.3: *Դիցուք M_1 և $M_2 \subseteq R^n$: Կասենք, որ այդ բազմությունները **խիստ** են անջատվում հիպերհարթությամբ, եթե գոյություն ունեն այնպիսի p վեկտոր և $\varepsilon > 0$ թիվ, որ*

$$(p, x) \leq (p, y) - \varepsilon \;\; \forall x \in M_1, \; \forall y \in M_2:$$

Թեորեմ 4.1.2: *Դիցուք M_1 և $M_2 \subseteq R^n$ բազմությունները **խիստ** են անջատվում հիպերհարթությամբ: Այդ դեպքում նրանք իրարից գտնվում են դրական հեռավորության վրա: Այսինքն՝*

$$d(M_1, M_2) = \inf_{x \in M_1, \, y \in M_2} \|x - y\| > 0:$$

Գծ.4.4-ում տրված է բազմությունների խիստ անջատման երկրաչափական մեկնաբանությունը հարթության վրա:

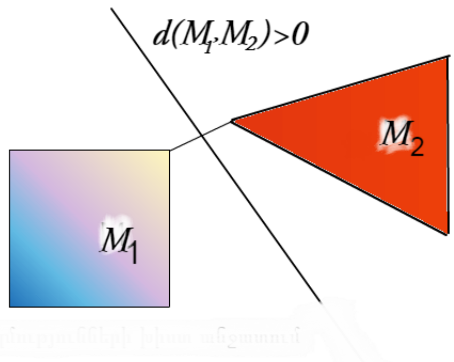
► Դիցուք M_1, M_2 բազմությունները խիստ են անջատվում հիպերհարթությամբ: Այդ դեպքում գոյություն ունեն այնպիսի p, ($\|p\| = 1$) վեկտոր և $\varepsilon > 0$ թիվ, որ

$$(p, x) \leq (p, y) - \varepsilon \;\; \forall x \in M_1, \; \forall y \in M_2:$$

Այստեղից կստանանք

$$\varepsilon = \inf_{y \in M_2} (p, y) - \sup_{x \in M_1} (p, x) = \inf_{x \in M_1, \, y \in M_2} (p, y - x) \leq$$

$$\leq \inf_{x \in M_1, \, y \in M_2} \|p\| \|x - y\| = d(M_1, M_2): \blacksquare$$



Գծ. 4.4: Ուռուցիկ բազմությունների խիստ անջատումը հիպերհարթությամբ

Հետագայում Լագրանժի անորոշ գործակիցների մեթոդի հիմնավոր շարադրման համար կարևոր է հետևյալ պնդումը, որը բերում ենք առանց ապացույցի (տես, օրինակ՝ [24], Թեորեմ 2.1-ի հետևանք 1, էջ 71):

Թեորեմ 4.1.3: *Դիցուք $M_1 \subset R^n$ փակ ուռուցիկ բազմություն է, իսկ $M_2 \subset R^n$ ուռուցիկ կոմպակտ է և*

$$M_1 \cap M_2 = \emptyset:$$

Այդ դեպքում նրանք խիստ անջատվում են հիպերհարթությամբ:

ԽՆԴԻՐՆԵՐ

1. Գրել այն հիպերհարթության հավասարումը, որը անջատում է $(-1, 2, 1, -3)$ կետը $M \subseteq R^4$ բազմությունից, որը տրվում է անհավասարությունների հետևյալ համակարգով.

$$\begin{cases} 5x_1 + x_2 - 5x_3 - 3x_4 \leq 1, \\ -3x_2 - 2x_2 + 5x_3 + x_4 \leq 2, \\ 3x_2 + x_3 + 2x_4 \leq 0, \\ x_1 + x_2 + 3x_3 - x_4 \leq 9: \end{cases}$$

2. Ապացուցել հետևյալ պնդումը: Որպեսզի $M \subseteq R^n$ փակ բազմությունը լինի ուռուցիկ, անհրաժեշտ է և բավարար, որ ցանկացած $b \notin M$ կետ խիստ անջատվի հիպերհարթությամբ M բազմությունից:

3. Դիցուք ուռուցիկ բազմության տրված կետով կարելի է տանել երկու հենման հիպերհարթություն: Ապացուցել, որ այդ կետով անցնում են անթիվ բազմության հենման հիպերհարթություններ:

4. Դիցուք $M_1, M_2 \subset R^n$ այնպիսի ուռուցիկ բազմություններ են, որ

$$intM_2 \neq \emptyset \text{ և } M_1 \cap intM_2 = \emptyset:$$

Ապացուցել, որ M_1 և M_2 բազմությունները անջատվում են հիպերհարթությամբ:

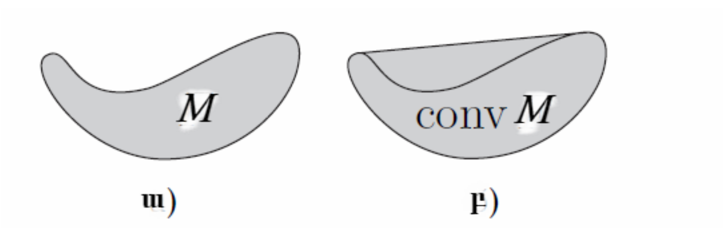
4.2 Կարաթեոդորիի թեորեմը

Սահմանում 4.2.1: *Դիցուք M-ը R^n-ի ենթաբազմություն է: Այդ բազմության **ուռուցիկ թաղանթ** կոչվում է հետևյալ բազմությունը.*

$$convM \equiv \{y \in R^n / y = \sum_{i=1}^{k} \alpha_i x^i,\ x^i \in M,$$

$$\sum_{i=1}^{k} \alpha_i = 1,\ \alpha_i \geq 0,\ i \in [1:k],\ k = 1, 2, ...\}$$

(տես գծ 4.5):



Գծ. 4.5: Բազմության ուռուցիկ թաղանթի օրինակ

Լեմմ 4.2.1: *Դիցուք ոչ զրոյական $x \in R^n$ վեկտորը ներկայացվում է $X = \{x^1, x^2, \dots\ , x^m\}$ համախմբի վեկտորների գծային ոչ բացասական կոմբինացիայով (գծային կոմբինացիայի գործակիցները ոչ բացասական թվեր են): Այդ դեպքում x-ը ներկայացվում է նաև այդ համակարգի գծորեն անկախ մի ենթահամակարգի վեկտորների գծային ոչ բացասական կոմբինացիայի միջոցով:*

► Դիցուք

$$x = \sum_{i=1}^{m} \alpha_i x^i, \ \alpha_i \geq 0, \ i \in [1:m]:$$

Եթե X համախմբի վեկտորները գծորեն անկախ են, ապա պնդումն ապացուցված է: Դիցուք այժմ X համախմբի վեկտորները գծորեն կախված են: Այդ դեպքում գոյություն ունեն այնպիսի $\lambda_1, \ \lambda_2, ..., \lambda_m$ թվեր, որոնցից գոնե մեկը զրո չէ, որ

$$x = \sum_{i=1}^{m} \lambda_i x^i:$$

Կարող ենք ենթադրել, որ այդ գործակիցներից որևէ մեկը դրական է: Նշանակենք

$$\alpha_0 \equiv \min_{\lambda_i > 0} \frac{\alpha_i}{\lambda_i} = \frac{\alpha_s}{\lambda_s}:$$

Ունենք

$$x = \sum_{i=1}^{m} \alpha_i x^i - \alpha_0 \sum_{i=1}^{m} \lambda_i x^i = \sum_{i=1}^{m} (\alpha_i - \alpha_0 \lambda_i) x^i:$$

Ակնհայտ է, որ

$$\alpha_s - \alpha_0 \lambda_s = 0 \text{ և } \alpha_i - \alpha_0 \lambda_i \geq 0 \ \forall i:$$

Այսպեղից հետևում է, որ x վեկտորը ներկացվում է X-ի ինչ-որ վեկտորների ոչ բացասական կոմբինացիայով, որոնց քանակը փոքր է m-ից: Այսպես շարունակելով՝ կհանգենք գծորեն անկախ համակարգի: ■

Թեորեմ 4.2.1: *Դիցուք $x \in R^n$ վեկտորը ներկայացվում է հետևյալ տեսքով.*

$$x = \sum_{i=1}^{m} \lambda_i x^i,\ \lambda_i \geq 0,\ i \in [1:m],\ \ \sum_{i=1}^{m} \lambda_i = 1,\ m > n: \tag{4.2.1}$$

Այդ դեպքում $X = \{x^1, ..., x^m\}$ համախմբի մեջ գոյություն ունի այնպիսի մի $\{x^{i_1}, x^{i_2}, ..., x^{i_{n+1}}\}$ ենթահամախումբ և ոչ բացասական $\alpha_{i_1} \geq 0,\ \alpha_{i_2} \geq 0, ...\ \alpha_{i_{n+1}} \geq 0$ թվեր, որ $\sum_{j=1}^{n+1} \alpha_{i_j} = 1$ և

$$x = \sum_{j=1}^{n+1} \alpha_{i_j} x^{i_j}:$$

► Դիտարկենք $(x, 1) \in R^{n+1}$ վեկտորը: Ըստ (4.2.1)-ի՝ ունենք

$$(x, 1) = \sum_{i=1}^{m} \lambda_i (x^i, 1) :$$

Համաձայն **Լեմմ 4.2.1**-ի գոյություն ունեն վեկտորների գծորեն անկախ այնպիսի $\{x^{i_1}, x^{i_2}, ..., x^{i_k}\}$ համախումբ և ոչ բացասական թվեր $\alpha_{i_1}, \alpha_2, , ..., \alpha_{i_k}$, որ

$$(x, 1) = \sum_{j=1}^{k} \alpha_{i_j} (x^{i_j}, 1):$$

Այսինքն՝

$$x=\sum_{j=1}^{k}\alpha_{i_j}x^{i_j},\ \alpha_{i_1}\geq 0,\ \alpha_{i_2}\geq 0,...,\alpha_{i_k}\geq 0,\ \sum_{j=1}^{k}\alpha_{i_j}=1: \tag{4.2.2}$$

Բայց, քանի որ $n+1$ չափանի պարածությունում գծորեն անկախ համակարգի վեկտորների քանակը $n+1$-ից ավել չէ, ապա $k\leq n+1$: Եթե $k=n+1$, ապա թեորեմն ապացուցված է: Եթե $k<n+1$, ապա ավելացնելով զրոյական անդամներ (4.2.2) գումարի անդամների քանակը դարձնում ենք $n+1$: ■

Հետևանք 4.2.1 *(Կարաթեոդորի թեորեմ): Դիցուք M-ը R^n-ի ենթաբազմություն է: Այդ դեպքում՝*

$$convM=\{y\in R^n/y=\sum_{i=1}^{n+1}\alpha_i x^i,\ x^i\in M,$$

$$\sum_{i=1}^{n+1}\alpha_i=1,\ \alpha_i\geq 0,\ i\in[1:n+1]\}:$$

ԽՆԴԻՐՆԵՐ

1. Գտնել հետևյալ բազմությունների ուռուցիկ թաղանթ–ները:

ա) $M=\{(x_1,x_2)\in R^2/x_1x_2=1\}$;

բ) $M=\{(x_1,x_2)\in R^2/x_2=\exp(-x_1)\}$;

գ) $M=\{(x_1,x_2)\in R^2/x_1,\ \ x_2\in[0,1]\}\cup\{(x_1,x_2)\in$ $\in R^2/x_2=x_1\}$:

2*. Ապացուցել, որ բաց բազմության ուռուցիկ թաղանթը բաց բազմություն է:

Ցուցում: Օգտվել այն փաստից, որ եթե բազմությունը ուռուցիկ է, ապա նրա ներքին կետերի բազմությունը նույնպես ուռուցիկ է:

3. Ապացուցել, որ կոմպակտ բազմության ուռուցիկ թաղանթը կոմպակտ է:

Ցուցում: Վերցնել հաջորդականություն բազմության ուռուցիկ թաղանթից և այն ներկայացնել Կարաթեոդորի թեորեմի օգնությամբ: Այնուհետև անցնել սահմանի այդ ներկայացման մեջ:

4. Արդյո՞ք անջատվում է $(1,-1,0)$ կետը

$$M = conv\{(-1,1,2),\ (2,-1,-3),$$

$$(-2,3,-1),\ (-5,-1,3)\}$$

բազմությունից հիպերհարթությամբ:

5. Ապացուցել, որ $convM$-ը մինիմալ այն ուռուցիկ բազմությունն է, որը պարունակում է M-ը:

6. Ապացուցել հետևյալ պնդումը: Որպեսզի $M \subseteq R^n$ բազմությունը լինի ուռուցիկ, անհրաժեշտ է և բավարար, որ

$$convM = M:$$

7. Դիցուք $M = conv\{x^1, x^2, ..., x^m\}$, իսկ $f(x)$-ը ուռուցիկ ֆունկցիա է, որոշված M բազմության վրա: Ապացուցել, որ

$$\max_{x \in M} f(x) = \max_{i \in [1:m]} f(x^i):$$

Ցուցում: Օգտվել Յենսենի անհավասարությունից:

8. Գտնել $f(x_1, x_2) = 25(x_1 - 2)^2 + (x_2 - 2)^2$ ֆունկցիայի մեծագույն արժեքը հետևյալ բազմության վրա.

$$\begin{cases} x_1 + x_2 \geq 2, \ x_1 - x_2 \geq -2, \\ x_1 + x_2 \leq 6, \ x_1 - 3x_2 \leq 2, \\ x_1, \ x_2 \geq 0 \qquad\qquad\qquad : \end{cases}$$

4.3 Հելլիի թեորեմը

Թեորեմ 4.3.1 *(Ռադոն)*: *Դիցուք տրված է վեկտորների* $X = \{x^1, x^2, ..., \ x^k\}$ $(k \geq n+2)$ *համախումբը* R^n*-ում: Այդ դեպքում այդ բազմությունը կարելի է տրոհել այնպիսի երկու ենթաբազմությունների՝* Y *և* Z*, որ*

$$convY \cap convZ \neq \emptyset:$$

► Դիտարկենք վեկտորների

$$\{x^2 - x^1, x^3 - x^1, ..., x^k - x^1\}$$

համախումբը: Քանի որ այս բազմության վեկտորների քանակը մեծ կամ հավասար է $n+1$, ապա նրանք գծորեն կախված են: Հետևաբար գոյություն ունեն այնպիսի $\alpha_2, \alpha_3, ..., \alpha_k$ թվեր, որոնցից գոնե մեկը զրո չէ, որ

$$\sum_{i=2}^{k} \alpha_i(x^i - x^1) = 0 \Rightarrow \sum_{i=1}^{k} \alpha_i x^i = 0, \text{ որտեղ } \alpha_1 = -\sum_{i=2}^{k} \alpha_i:$$

Այսպիսով, գոյություն ունեն այնպիսի $\alpha_1, ..., \ \alpha_k$ թվեր, որոնցից գոնե մեկը հավասար չէ զրոյի, որ

$$\sum_{i=1}^{k} \alpha_i x^i = 0, \ \sum_{i=1}^{k} \alpha_i = 0 : \qquad (4.3.1)$$

Այն α_i գործակիցները, որոնք հավասար են զրոյի դեն գցենք և ընդհանրությունը չխախտելով ենթադրենք, որ

$$\alpha_1 > 0, ..., \ \alpha_{k'} > 0, \ \alpha_{k'+1} < 0, ..., \alpha_k < 0:$$

Նշանակենք

$$\nu = \sum_{i=1}^{k} \alpha_i = - \sum_{i=k'+1}^{k} \alpha_i:$$

(4.3.1)-ից կստանանք

$$\sum_{i=1}^{k'} \frac{\alpha_i}{\nu} x^i = \sum_{i=k'+1}^{k} -\frac{\alpha_i}{\nu} x^i : \qquad (4.3.2)$$

Նշանակելով

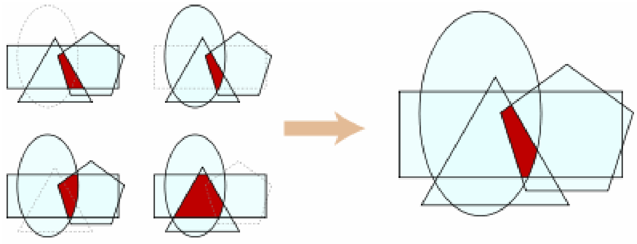
$$Y = \{x^1, ..., x^{k'}\}, \quad Z = \{x^{k'+1}, ..., x^k\}$$

(4.3.2)-ից՝ կստանանք

$$x \equiv \sum_{i=1}^{k'} \frac{\alpha_i}{\nu} x^i = \sum_{i=k'+1}^{k} -\frac{\alpha_i}{\nu} x^i \in convX \cap convY: \blacksquare$$

Թեորեմ 4.3.2 *(Հելլի)*: *Եթե $\{A_\alpha\}$-ն R^n տարածության կոմպակտ ուռուցիկ ենթաբազմությունների այնպիսի ընտանիք է, որ նրա ցանկացած $n+1$ քանակով բազմությունների հատումը դատարկ չէ, ապա*

$$\bigcap_\alpha A_\alpha \neq \emptyset \textit{ (տես գծ. 4.6)} :$$



Գծ. 4.6: Հելլիի թեորեմի երկրաչափական մեկնաբանությունը հարթության վրա

► Քանի որ A_α բազմությունները կոմպակտ են, ապա բավական է ցույց տալ, որ այդ ընտանիքի ցանկացած վերջավոր քանակով բազմությունների հատումը դատարկ չէ: Ապացույցը կատարենք ինդուկցիայով ըստ բազմությունների քանակի: Դիցուք $A_1, A_2, ..., A_k$ $(k > n+1)$ բազմություններ են ընտանիքից և այդ բազմություններից ցանկացած $k-1$-ի հատումը դատարկ չէ: Ապացուցենք, որ

$$\bigcap_{i=1}^{k} A_i \neq \emptyset :$$

Նշանակենք

$$D_i = A_1 \cap A_2 \cap ... \cap A_{i-1} \cap A_{i+1} ... \cap A_k, \quad i \in [1:k]: \quad (4.3.3)$$

Ըստ ենթադրության D_i, $i \in [1 : k]$, բազմությունները դատարկ չեն: Ընտրենք կամայական $x^i \in D_i$ էլեմենտներ և ձևավորենք $X = \{x^1, x^2, ..., x^k\}$ համախումբը: Ըստ **Ռադոնի թեորեմի** այդ համախումբը կարելի է տրոհել այնպիսի երկու մասերի, որ $X = Y \bigcup Z$, $Y \bigcap Z = \emptyset$ և

$$convY \bigcap convZ \neq \emptyset:$$

Ենթադրենք

$$Y = \{x^1, x^2, ..., x^{k'}\},\ Z = \{x^{k'+1}, ..., x^k\}:$$

Ցույց տանք, որ եթե

$$x \in convY \bigcap convZ,$$

ապա $x \in \bigcap_{i=1}^k A_i$: Ունենք

$$x = \sum_{i=1}^{k'} \lambda_i x^i,\ \ \sum_{i=1}^{k'} \lambda_i = 1, \quad \lambda_i \geq 0,\, i \in [1 : k']:$$

Այսպեղից քանի որ

$$x^i \in \bigcap_{j=k'+1}^{k} A_j,\ \forall i \leq k'$$

և $A_j,\ j \in [1 : k]$, բազմությունները ուռուցիկ են, ապա $x \in \bigcap_{j=k'+1}^k A_j$: Մյուս կողմից, քանի որ

$$x = \sum_{j=k'+1}^{k} \lambda_j x^j,$$

ապա $x \in \bigcap_{j=1}^{k'} A_j$: Այսպիսով, $x \in \bigcap_{j=1}^k A_j$: ■

ԽՆԴԻՐՆԵՐ

1. Դիցուք R^2 հարթության n հատ կետեր բավարարում են հետևյալ պայմանին. նրանցից ցանկացած երեքը գտնվում են միավոր շառավղով ինչ որ շրջանի ներսում: Ապացուցել, որ բոլոր կետերն են գտնվում միավոր շառավղով միևնույն շրջանի ներսում:
Ցուցում: Դիցուք $a^1, a^2, ..., a^n$-ը տրված կետեր են: Դիտարկել

$$A_i = \{x \in R^2 / \|x - a^i\| \leq 1\}, \ \ i \in [1:n]$$

շրջանների բազմությունը և այդ ընտանիքի նկատմամբ կիրառել Հելլիի թեորեմը:

2. Դիցուք $M \subset R^n$ ուռուցիկ կոմպակտ բազմությունը ծածկված է բաց կիսատարածությունների ընտանիքով: Ապացուցել, որ այդ ընտանիքից կարելի է ընտրել $n+1$ հատ կիսատարածություններ, որոնք նույնպես ծածկում են M-ը:

3. Դիցուք $M \subset R^2$-ը d տրամագծով կոմպակտ բազմություն է: Ապացուցել, որ գոյություն ունի $d/\sqrt{3}$ շառավղով շրջան, որը պարունակում է M-ը:
Ցուցում: Նախ ցույց տալ, որ այդ բազմության ցանկացած երեք կետերի համար գոյություն ունի $d/\sqrt{3}$ շառավղով շրջան, որը պարունակում է այդ կետերը: Այնուհետև այդ շրջանների նկատմամբ կիրառել Հելլիի թեորեմը:

Նշենք, որ ընթերցողը Հելլիի թեորեմի բազմաթիվ կիրա– ռությունների հետ կարող է ծանոթանալ [26]-ում:

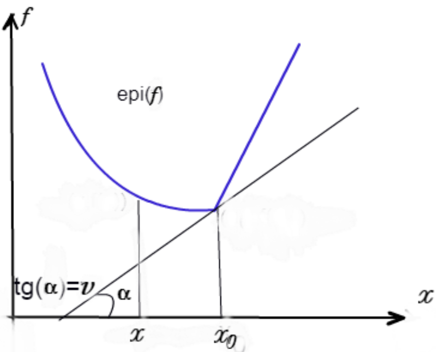
4.4 Ուռուցիկ ֆունկցիայի սուբդիֆերենցիալը

Սահմանում 4.4.1: *Դիցուք f-ը ուռուցիկ ֆունկցիա է՝ որոշված R^n-ի վրա: v վեկտորը կոչվում է սուբգրադիենտ x^0 կետում, եթե տեղի ունի*

$$f(x) - f(x^0) \geq (v, x - x^0)\ \forall x \in R^n \tag{4.4.1}$$

անհավասարությունը:
f-ի սուբգրադիենտների բազմությունը x^0 կետում կոչվում է ***սուբդիֆերենցիալ*** *և նշանակվում է $\partial f(x^0)$ սիմվոլով:*

Գծ.4.7-ում ցույց է տրված, որ մեկ փոփոխականի f ուռուցիկ ֆունկցիայի դեպքում v սուբգրադիենտը գրաֆիկի $(x_0, f(x_0))$ կետով տարված հենման $y = f(x_0) + (v, x - x_0)$ գծի և x առանցքի հետ կազմած անկյան տանգեսն է:



Գծ. 4.7: Սուբգրադիենտի երկրաչափական իմաստը

Թեորեմ 4.4.1: *$\partial f(x^0)$ բազմությունը ոչ դատարկ ուռուցիկ կոմպակտ է:*

► Նախ ցույց տանք $\partial f(x^0)$-ի ուռուցիկությունը:
Եթե $v^1 \in \partial f(x^0), v^2 \in \partial f(x^0)$, ապա

$$f(x) - f(x^0) \geq (\alpha v^1 + (1-\alpha)v^2, x - x^0) \ \forall x \in R^n, \ \forall \alpha \in [0,1],$$

որտեղից հետևում է $\partial f(x^0)$-ի ուռուցիկությունը:
$\partial f(x^0)$ բազմության փակությունը ակնհայտ է: Ապացուցենք, որ այն դատարկ չէ: Դիտարկենք f ֆունկցիայի վերգրաֆիկը.

$$epi(f) = \{(\beta, x) \in R^{n+1} / \beta \geq f(x)\}:$$

Քանի որ այս բազմությունը փակ է և ուռուցիկ, ապա նրա եզրային $(f(x^0), x^0)$ կետից կարելի է տանել հենման հիպերհարթություն: Դա նշանակում է, որ գոյություն ունի այնպիսի ոչ զրոյական $(c, v) \in R^{n+1}$ վեկտոր, որ

$$c\beta + (v, x) \geq cf(x^0) + (v, x^0) \ \forall (\beta, x) \in epi(f): \qquad (4.4.2)$$

Եթե $c = 0$, ապա (4.4.2)-ից ստանում ենք

$$(v, x - x^0) \geq 0 \ \forall x \in R^n$$

անհավասարությունը: Այստեղից հետևում է, որ $v = 0$, որը հակասություն է, քանի որ $(c, v) \neq 0$: Ցույց տանք, որ $c > 0$: (4.4.2) անհավասարությունը $x = x^0$ դեպքում ունի

$$c(\beta - f(x^0)) \geq 0$$

տեսքը, որտեղից անմիջականորեն հետևում է, որ $c > 0$: Այժմ (4.4.2) անհավասարության երկու մասերը բաժանենք $c > 0$ թվի վրա և այնուհետև տեղադրելով $\beta = f(x)$ կստանանք

$$f(x) - f(x^0) \geq (-\frac{v}{c}, x - x^0),$$

որտեղից հետևում է, որ

$$-\frac{v}{c} \in \partial f(x^0) :$$

Ապացուցենք, որ $\partial f(x^0)$ բազմությունը սահմանափակ է: Ենթադրենք հակառակը: Այդ դեպքում գոյություն կունենա այնպիսի $\{v^k\}$ $(v^k \in \partial f(x^0))$ հաջորդականություն, որ $\|v^k\| \to \infty$: (4.4.1) անհավասարությունում տեղադրելով $x = x^0 + v^k/\|v^k\|$, կստանանք

$$f(x) - f(x^0) \geq \|v^k\|: \tag{4.4.3}$$

Քանի որ f-ը անընդհատ է, ապա (4.4.3) անհավասարության ձախ մասը սահմանափակ է, իսկ աջ մասը ձգտում է անվերջության, երբ $k \to \infty$, ինչը հակասություն է: ■

Ֆունկցիայի մինիմիզացիայի ալգորիթմների մշակման համար կարևոր է նկարագրել և հաշվել ֆունկցիայի նվազման ուղղությունները յուրաքանչյուր կետում: Դիֆերենցելիության դեպքում այդ հարցը մեխանիկորեն լուծվում է՝ նվազման ուղղությունը հակագրադիենտի ուղղությունն է: Սակայն երբ ֆունկցիան դիֆերենցելի չէ այդ խնդիրը բարդ է: Բայց ուռուցիկ որոշ դասի ֆունկցիաների համար սուբգրադիենտի օգնությամբ հնարավոր է նկարագրել ֆունկցիաների նվազման ուղղությունները և կառուցել մինիմիզացիայի ալգորիթմներ: Սպորև բերված պնդումները նկարագրում են այդ ֆունկցիաները և նրանց նվազման ուղղությունները: Ճիշտ են հետևյալ պնդումները:

Թեորեմ 4.4.2 [12] *(Թեորեմ 2.1, էջ 71): Դիցուք* $f_1(x),\ f_2(x), ..., f_k(x)$ *դիֆերենցելի ուռուցիկ ֆունկցիաներ են: Նշանակենք*

$$f(x) \equiv \max_{i \in [1:k]} f_i(x):$$

Այդ դեպքում

$$\partial f(x) = conv\{f_i'(x),\ i \in I(x)\},$$

$$\text{որտեղ}\ \ I(x) = \{i \in [1:k]/f_i(x) = f(x)\}։$$

Թեորեմ 4.4.3 [12] *(Թեորեմ 3.3, էջ 85)։ Դիցուք տեղի ունեն **Թեորեմ 4.4.2-ի** պայմանները և*

$$0 \notin \partial f(x)։$$

Այդ դեպքում

$$h = -\frac{\Pi_{\partial f(x)}(0)}{\|\Pi_{\partial f(x)}(0)\|}$$

վեկտորը f ֆունկցիայի նվազման ուղղությունն է x կետում։
Վերը նշված թեորեմները հնարավորություն են տալիս ամենաարագ վայրէջքի մեթոդով որոշ դեպքերում գտնելու ուռուցիկ ոչ դիֆերենցելի ֆունկցիայի մինիմումի կետերը։ Մեկնաբանենք դա օրինակի միջոցով։

Օրինակ։ Դիցուք

$$f(x_1, x_2) \equiv \max_{i \in [1:3]} f_i(x_1, x_2),$$

որտեղ

$$f_1(x_1, x_1) = 2x_1 + x_2,\ f_2(x_1, x_2) = -x_1 - 2x_2,$$

$$f_3(x_1, x_2) = -x_1 + x_2 - 3։$$

Պետք է գտնել f-ի մինիմումի կետը R^2-ի վրա։ Որպես սկզբնական մոտավորություն վերցնենք $x^0 = (0, 0)$ կետը։ Քանի որ

$$f_1(x^0) = f_2(x^0) = 0,\ f_3(x^0) = -3,$$

ապա $I(x^0) = \{1, 2\}$, և հետևաբար ըստ **Թեորեմ 4.4.2 -ի** կունենանք՝

$$\partial f(x^0) = conv\{f_1'(0,0), f_2'(0,0)\} = conv\{(2,1),\ (-1,-2)\}:$$

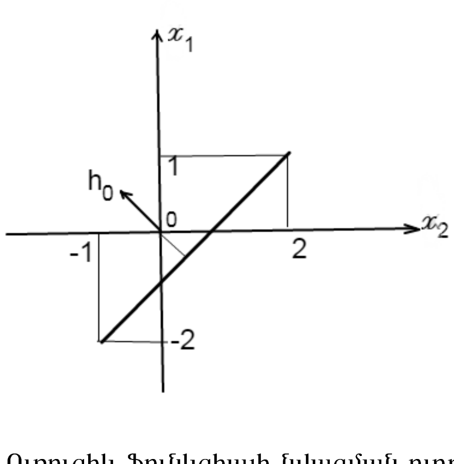
Պարզ է, որ

$$0 \notin \partial f(x^0),$$

ուստի, ըստ **Թեորեմ 4.4.3-ի՝**

$$h^0 = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$$

վեկտորը կլինի f-ի նվազման ուղղությունը x^0 կետում (տես գծ. 4.8):



Գծ. 4.8: Ոււռուցիկ ֆունկցիայի նվազման ուղղությունը

Այժմ ամենաարագ վայրէջքի մեթոդով գտնենք α_0 քայլի երկարությունը և կառուցենք x^1 կետը հետևյալ բանաձևով.

$$x^1 = x^0 + \alpha_0 h^0:$$

Ունենք

$$g(\alpha) \equiv f(x^0 + \alpha h^0) = max\{-\frac{\sqrt{2}}{2}\alpha, -\frac{\sqrt{2}}{2}\alpha, \sqrt{2}\alpha - 3\} =$$

$$= \frac{\sqrt{2}}{2}\, max\{-\alpha, 2\alpha - 3\sqrt{2}\}:$$

Այսպեղից

$$\alpha_0 = arg \min_{\alpha \in (0, +\infty)} f(x^0 + \alpha h^0) = \sqrt{2}:$$

Հեփևաբար $x^1 = x^0 + \alpha_0 h^0 = (-1, 1)$:
Քանի որ

$$f_1(x^1) = f_2(x^1) = f_3(x^1) = -1,$$

ապա $I(x^1) = \{1, 2, 3\}$:

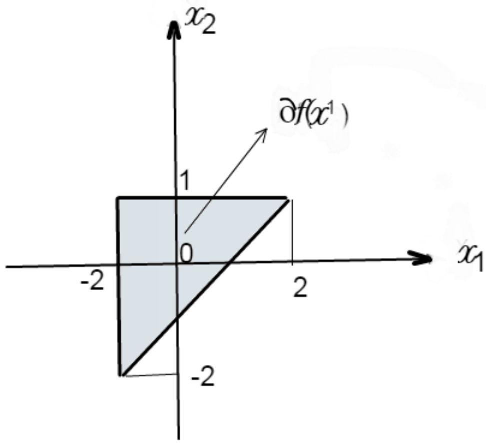
Հեփևաբար

$$\partial f(x^1) = conv\{f_1'(x^1),\ f_2'(x^1), f_3'(x^1)\} =$$

$$= conv\{(2, 1)\ (-1, -2)\ (-1, -1)\}:$$

Հեշփ է նկափել, որ $0 \in \partial f(x^1)$ (փես գծ 4.9):

Այսպեղից հեփևում է, որ x^1 վեկփորը f ուռուցիկ ֆունկցիայի մինիմումի կեփ է:



Գծ. 4.9: Ուռուցիկ ֆունկցիայի սուբդիֆերենցիալը

Ինչպես երևում է **Թեորեմ 4.4.3-ից** ուռուցիկ ֆունկցիայի նվազման ուղղությունները գտնելու համար կարևոր խնդիր է որոշել 0 կետի պրոյեկցիան ուռուցիկ կոմպակտի վրա: Նկարագրենք մի իտերացիոն ալգորիթմ, որի միջոցով ցանկացած ճշտությամբ կարելի է գտնել 0 կետի պրոյեկցիան ուռուցիկ բազմանիստի վրա: Դիցուք տրված է n չափանի վեկտորների $X = \{x^1,\ x^2, ..., x^m\}$ համախումբը: Պետք է գտնել 0 կետի պրոյեկցիան $M = convX$ բազմության վրա, այսինքն՝ $\Pi_M(0)$-ն:

Կառուցում ենք $\{v^k\}$ հաջորդականությունը հետևյալ կերպ:

- Որպես v^0 սկզբնական մոտավորություն վերցնում ենք այն վեկտորը X համախմբից, որը բավարարում է հետևյալ պայմանին.

$$(v^0, v^0) = \min_{i \in [1:m]} (x^i, x^i):$$

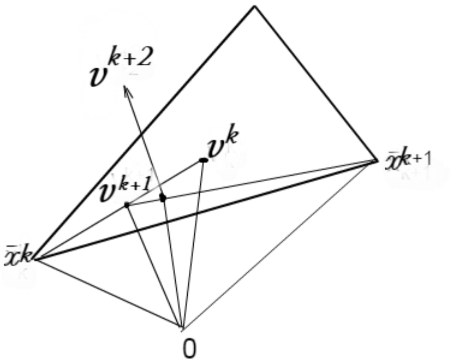
- Ենթադրենք արդեն ունենք $v^k \in M$ վեկտորը: Նկարագրենք v^{k+1} կառուցման ալգորիթմը: Ընտրում ենք X համախմբից այն $\overline{x}^k$ վեկտորը, որը բավարարում է
$$(\overline{x}^k, v^k) = \min_{i \in [1:m]} (v^k, x^i)$$
հավասարությանը:
Այնուհետև հաշվում ենք
$$\delta(v^k) \equiv (v^k - \overline{x}^k, v^k), \;\; t_k = \frac{\delta(v^k)}{\|\overline{x}^k - v^k\|}$$
մեծությունները:

- v^{k+1} վեկտորը կառուցում ենք հետևյալ ռեկուրենտ առնչությամբ.
$$v^{k+1} = v^k + t_k(\overline{x}^k - v^k):$$

Ճշմարիտ է հետևյալ պնդումը, որը բերում ենք առանց ապացույցի (տես, օրինակ՝ [12], Թեորեմ 2, էջ 340):

Թեորեմ 4.4.4: *Դիցուք $\{v^k\}$ հաջորդականությունը կառուցվել է վերը նշված ալգորիթմով:*
Այդ դեպքում $v^k \to \Pi_M(0)$, երբ $k \to \infty$:

Առաջադրանք: Կատարել վերը նշված ալգորիթմի ծրագրային իրականացումը C^{++} լեզվով:
Գտնել $\Pi_M(0)$ վեկտորը, եթե
$$M = conv\{(1,1,1), \; (2,2,2), \dot{(}1, 1.5, 1), \; (2,3,4)\}:$$



Գծ. 4.10: $\{v^k\}$ հաջորդականության կառուցման ընթացակարգը

ԽՆԴԻՐՆԵՐ

1. Դիցուք f ուռուցիկ ֆունկցիան դիֆերենցելի է x^0 կետում: Ապացուցել, որ

$$\partial f(x^0) = \{f'(x^0)\}:$$

2. Դիցուք $f(x) = \|x\|$: Ցույց տալ, որ

$$\partial f(0) = B_1(0):$$

3. Հաշվել հետևյալ ուռուցիկ ֆունկցիաների սուբդիֆերենցիալները:

ա) $f(x) = |x-1| + |x+1|,$

բ) $f(x) = max\{4x+1, x-2\}$,

գ) $f(x_1, x_2) = |x_1 - 1/2|x_2| + x_2|$,

դ) $f(x_1, x_2) = |x_1 + x_2|$,

ե) $f(x_1, x_2) = ||x_1| + x_2 - 1|$:

4. Ապացուցել հետևյալ պնդումը: Որպեսզի x^0 կետը լինի f ուռուցիկ ֆունկցիայի մինիմումի կետ R^n-ի վրա, անհրաժեշտ է և բավարար, որ

$$0 \in \partial f(x^0):$$

5. c պարամետրերի բոլոր արժեքների դեպքում լուծել հետևյալ խնդիրները.

ա) $\|x\| - (c, x) \to min, \ x \geq 0, \ x \in R^n$,

բ) $\frac{1}{2}\|x\|^2 + \|x - c\| \to min, \ x \in R^n$,

6. $f(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 + 3|x_1 + x_2 - 2| \to min$:

Լուծում: Որպեսզի (x_1, x_2) կետը լինի f-ի մինիմումի կետ, անհրաժեշտ է և բավարար, որ գոյություն ունենա այնպիսի $\mathbf{v} \in \partial g(x_1, x_2)$ վեկտոր, որ

$$0 \in \partial f(x_1, x_2) \Leftrightarrow (2x_1 + x_2, 2x_2 + x_1) + 3\mathbf{v} = 0, \qquad (4.4.3)$$

որտեղ $g(x_1, x_2) \equiv |x_1 + x_2 - 2|$: Դժվար չէ տեսնել, որ

$$\partial g(x_1, x_2) = \begin{cases} (1,1), & \text{եթե } x_1 + x_2 > 2 \\ (-1,-1), & \text{եթե } x_1 + x_2 < 2 \\ conv\{(1,1), (-1,-1)\}, & \text{եթե } x_1 + x_2 = 2: \end{cases}$$

Եթե $x_1 + x_2 > 2$, ապա (4.4.3) պայմանից կստանանք

$$\begin{cases} 2x_1 + x_2 + 3 = 0, \\ 2x_2 + x_1 + 3 = 0 \end{cases} \Rightarrow x_1 + x_2 = -2,$$

այսինքն՝ համակարգը համատեղելի չէ:
Եթե $x_1 + x_2 < 2$, ապա համակարգը նույնպես լուծում չունի:
Եթե $x_1 + x_2 = 2$, ապա

$$\begin{cases} 2x_1 + x_2 + 3\alpha = 0, \\ 2x_2 + x_1 + 3\alpha = 0, \\ \alpha \in [-1, 1]: \end{cases}$$

Այստեղից կստանանք $\alpha = -1, x_1 = x_2 = 1$, այսինքն՝ $(1, 1)$ կետը խնդրի լուծումն է:

7. $x_1^2 + x_2^2 + 2max(x_1, x_2) \to min$:

8. $x_1^2 + x_2^2 + 2\sqrt{(x_1 - 1)^2 + (x_2 - 2)^2} \to min$:

9. $x_1^2 + x_2^2 + 4|x_1 + x_2 - 1| \to min$:

10. Ամենաարագ վայրէջքի մեթոդով գտնել

$$f(x_1, x_2) \equiv \max_{i \in [1:3]} f_i(x_1, x_2)$$

ֆունկցիայի մինիմումի կետը: Կատարել իտերացիայի երկու քայլ՝ որպես սկզբնական մոտավորություն վերցնելով $(-2.5, 0)$ կետը: Այստեղ

$$f_1(x_1, x_1) = x_1 + x_2,\ f_2(x_1, x_2) = -2x_1 - x_2,$$

$$f_3(x_1, x_2) = -x_1 + x_2 - 5:$$

4.5 Կուն-Տակկերի թեորեմը

Դիցուք $f_i(x), \ i \in [0 : m]$ ֆունկցիաները ուռուցիկ են R^n-ի վրա: Դիտարկենք $f_0(x)$ ֆունկցիայի մինիմիզացիայի խնդիրը $M \equiv \{x \in R^n / f_i(x) \leq 0, \ i \in [1 : m]\}$ բազմության վրա.

$$f_0(x) \to min, \ x \in M): \tag{4.5.1}$$

Հեշտ է ցույց տալ, որ M-ը ուռուցիկ բազմություն է: (3.5.1) խնդիրը կոչվում է **ուռուցիկ ծրագրավորման** խնդիր: $f_0(x)$ ֆունկցիայի մինիմումի կետերը M բազմության վրա կոչվում են (3.5.1) խնդրի լուծումներ: Կազմենք **Լագրանժի** ֆունկցիան.

$$L(x, \lambda) = \lambda_0 f_0(x) + \lambda_1 f_1(x) + ... + \lambda_m f_m(x),$$

որտեղ $\lambda = (\lambda_0, \lambda_1, ..., \lambda_m) \in R^{m+1}$:

Թեորեմ 4.5.1 *(Կուն-Տակկեր: Անհրաժեշտությունը)*: *Դիցուք x^* կետը հանդիսանում է (4.5.1) խնդրի լուծում: Այդ դեպքում գոյություն ունեն այնպիսի $\lambda_0 \geq 0, \lambda_1 \geq 0, ..., \lambda_m \geq 0$ թվեր, որոնցից գոնե մեկը զրո չէ, որ*

1) $L(x, \lambda) \geq L(x^*, \lambda) \ \forall x \in R^n$,

2) $\lambda_i f_i(x^*) = 0, \ i \in [1 : m]$*:*

► Չսահմանափակելով ընդհանրությունը, կարելի է ենթադրել, որ $f_0(x^*) = 0$: Հակառակ դեպքում կդիտարկենք $\tilde{f}_0(x) = f_0(x) - f_0(x^*)$ ֆունկցիան, որը նույնպես ուռուցիկ է և $\tilde{f}_0(x^*) = 0$: Դիտարկենք հետևյալ բազմությունը.

$$C = \{(\mu_0, \mu_1, ..., \mu_m) \in R^{m+1} / \exists x \in R^n :$$

$$f_0(x) < \mu_0,\ f_i(x) \le \mu_i,\ i \in [1:m]\}:$$

Այս բազմությունը ուռուցիկ է: Այն անմիջապես հետևում է սահմանումից: C-ն դատարկ չէ: Իսկապես, քանի որ

$$f_0(x^*) < 1,\ f_i(x^*) \le 1,\ i \in [1:m],$$

ապա

$$(1, 1, ..., 1) \in C:$$

Ակնհայտ է նաև, որ $0 \notin C$, քանի որ հակառակ դեպքում գոյություն կունենար այնպիսի x կետ, որ

$$f_0(x) < 0,\ f_i(x) \le 0,\ i \in [1:m],$$

ինչը հակասություն է:
0 կետը անջատենք հիպերհարթությամբ C ուռուցիկ բազմությունից: Դա նշանակում է, որ գոյություն ունեն այնպիսի $\lambda_i,\ i \in [0:m]$ թվեր, որոնցից գոնե մեկը զրո չէ, որ

$$\sum_{i=0}^{m} \lambda_i \mu_i \ge 0\ \ \forall(\mu_0, \mu_1, ..., \mu_m) \in C : \tag{4.5.2}$$

Ցույց տանք, որ $\lambda_i \ge 0,\ i \in [0:m]$: Ենթադրենք, որ ինչ որ i_0 ինդեքսի համար տեղի ունի $\lambda_{i_0} < 0$ անհավասարությունը: Ակնհայտ է, որ ցանկացած $\delta > 0$ և $\mu_{i_0} > 0$ թվի համար $\mu \equiv (\delta, 0, .., \mu_{i_0}, 0, ..0) \in C$: Տեղադրելով μ վեկտորի կոորդինատները (4.5.2) անհավասարությունում՝ կստանանք

$$\delta\lambda_0 + \mu_{i_0}\lambda_{i_0} \ge 0: \tag{4.5.3}$$

Այստեղ անցնելով սահմանի, երբ $\delta \to 0$ և $\mu_{i_0} \to +\infty$ կստանանք, որ (4.5.3) անհավասարության ձախ մասը ձգտում է $-\infty$, իսկ աջ մասը զրո է, որը հակասություն է:

Այժմ ապացուցենք թեորեմի երկրորդ պնդումը: (4.5.2) անհավասարությունում փեղադրելով

$$\mu_0 = \delta > 0, \mu_1 = 0, ..., \mu_i = f_i(x^*), \mu_{i+1} = 0, .., \mu_m = 0$$

թվերը՝ կսփանանք

$$\delta\lambda_0 + \lambda_i f_i(x^*) \geq 0:$$

Այսփեղ անցնելով սահմանի, երբ δ ձգփի զրոյի՝ կսփանանք

$$\lambda_i f_i(x^*) \geq 0 : \qquad (4.5.4)$$

Բայց, քանի որ $\lambda_i \geq 0$ և $f_i(x^*) \leq 0$, ապա (4.5.4)-ից հեփևում է, որ

$$\lambda_i f_i(x^*) = 0 : \qquad (4.5.5)$$

Այժմ ապացուցենք թեորեմի եզրակացության առաջին կեփը: Իրոք, ցանկացած $x \in R^n$-ի համար (4.5.3) անհավասարությունում փեղադրելով

$$\mu_0 = \delta + f_0(x), \mu_1 = f_1(x), ..., \mu_m = f_m(x)$$

կոորդինափներով μ վեկփորը, կսփանանք

$$\lambda_0(\delta + f_0(x)) + \lambda_1 f_1(x) + ... + \lambda_m f_m(x) \geq 0:$$

Անցնելով սահմանի, երբ $\delta \to 0$, կսփանանք՝

$$\lambda_0 f_0(x) + \lambda_1 f_1(x) + ... + \lambda_m f_m(x) \geq 0 \; \forall x \in R^n : \qquad (4.5.6)$$

(4.5.5)-(4.5.6)-ից հեփևում է որ

$$L(x, \lambda) = \sum_{i=0}^{m} \lambda_i f_i(x) \geq 0 =$$

$= \lambda_0 f_0(x^*)+\lambda_1 f_1(x^*)+...+\lambda_m f_m(x^*) = L(x^*,\lambda)\ \forall x \in R^n :$ ■

Անհրաժեշտության երկրորդ պայմանը կոչվում է պասիվ-ակտիվ սահմանափակումների պայման: Այս պայմանից անմիջապես հետևում է, որ այն սահմանափակումը, որի դեպքում $f_i(x^*) < 0$, ապա նրան համապատասխանող λ_i գործակիցը հավասար է զրոյի: Այդպիսի սահմանափակումները կոչվում են պասիվ, իսկ $f_i(x^*) = 0$ պայմանին բավարարող i-րդ սահմանափակումը կոչվում է ակտիվ:

Թեորեմ 4.5.2 *(Կուն-Տակկեր: Բավարարությունը)*: *Դիցուք* x^* *կետը բավարարում է* (4.5.1) *խնդրի բոլոր սահմանափակումներին.*

$$f_i(x^*) \le 0,\ i \in [1:m],$$

և գոյություն ունեն այնպիսի $\lambda_0 > 0,\ \lambda_1 \ge 0, \lambda_2 \ge \ge 0, ...,\ \lambda_m \ge 0$ *թվեր, որ տեղի ունեն թեորեմ 4.5.1-ի 1)-2) պայմանները:*
Այդ դեպքում x^* *վեկտորը* (4.5.1) *խնդրի լուծում է:*

▶ Չսահմանափակելով ընդհանրությունը՝ կարող ենք ենթադրել, որ $\lambda_0 = 1$: Այդ դեպքում, հաշվի առնելով 1) և 2) պայմանները, կունենանք՝

$$f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x) \ge f_0(x^*)+\sum_{i=1}^{m}\lambda_i f_i(x^*) = f_0(x^*)\ \forall x \in R^n : \tag{4.5.7}$$

Եթե x վեկտորը այնպիսին է, որ $f_i(x) \le 0\ \forall i \in [1:m]$, ապա (3.5.7) անհավասարությունից ստանում ենք

$$f_0(x) \ge f_0(x^*):$$ ■

Ռեգուլյարության պայմանը: Այս պայմանը (4.5.1) խնդրում հետևյալն է. գոյություն ունի այնպիսի $\bar{x}$ վեկտոր, որ

$$f_i(\bar{x}) < 0,\ i \in [1:m]:$$

Ցույց տանք, որ այս պայմանի դեպքում λ_0 գործակիցը կարելի է վերցնել հավասար մեկի: Իրոք, եթե $\lambda_0 = 0$, ապա Թեորեմ 4.5.1-ի 1) և 2) եզրակացություններից հետևում է, որ

$$\sum_{i=1}^{m} \lambda_i f_i(\bar{x}) \geq 0,$$

որը հակասություն է, քանի որ

$$\sum_{i=1}^{m} \lambda_i f_i(\bar{x}) < 0:$$

Նշենք, որ եթե $f_i(x)$, $i \in [1:m]$, ֆունկցիաները անընդհատ են, ապա ռեգուլյարության պայմանը ուռուցիկ ծրագրավորման խնդրում նույն է ինչ-որ M բազմությունը ունի ներքին կետ: ■

Դիֆերենցելի ֆունկցիաների դեպքում տանք Կուն-Տակկերի թեորեմի ևս մի ապացույց և երկրաչափական մեկնաբանություն: Դիտարկենք $f_0(x_1, x_2)$ ուռուցիկ դիֆերենցելի ֆունկցիայի մինիմումի խնդիրը $f_1(x_1, x_2) \leq 0$, $f_1(x_1, x_2) \leq 0$ սահմանափակումների դեպքում, որտեղ f_1-ը և f_2-ը նույնպես ուռուցիկ են և դիֆերենցելի: Ենթադրենք, որ $x*$-ը խնդրի լուծումն է և $f_1(x_1^*, x^*) = f_2(x_1^*, x_2^*) = 0$: Ենթադրենք նաև տեղի ունի ռեգուլյարության պայմանը և $f_0(x^*) = 0$: Դիտարկենք գծային անհավասարությունների հետևյալ համակարգը.

$$(f_0'(x^*), x - x^*) < 0,$$

$$(f_1'(x^*), x - x^*) < 0,$$

$$(f_2'(x^*), x - x^*) < 0:$$

Նախ, քանի որ տեղի ունի ռեգուլյարության պայմանը, ապա գոյություն ունի այնպիսի $\overline{x}$ վեկտոր, որ $f_i(\overline{x}) < 0, \ i = 1, 2$: Այստեղից, հաշվի առնելով նաև ուռուցիկ ֆունկցիայի հիմնական անհավասարությունը, կստանանք

$$(f_i'(x^*), \overline{x} - x^*) \leq f_i(\overline{x}) - f_i(x^*) = f_i(\overline{x}) < 0, \ i = 1, 2:$$

Այսպիսով, վերևում դիտարկված համակարգի երկրորդ և երրորդ անհավասարումներից կազմված համակարգը համատեղելի է: Նկատենք միաժամանակ, որ ընդհանուր համակարգը համատեղելի չէ:

Իրոք, եթե համակարգը լուծում ունենա, ապա բավականաչափ փոքր դրական α թվերի համար կունենանք

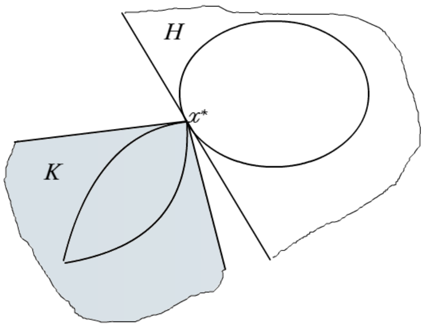
$$f_i(x^* + \alpha(x - x^*)) = f_i(x^*) + \alpha((f_i'(x^*), x - x^*) + \frac{o(\alpha)}{\alpha}) < 0,$$

ինչը հակասություն է, քանի որ x^*-ը f_0 ֆունկցիայի մինիմումի կետն է $f_i(x) \leq 0, \ i = 1, 2$ սահմանափակումների դեպքում: Նշանակենք

$$K = \{x \in R^2 / (f_1'(x^*), x - x^*) < 0, \ (f_2'(x^*), x - x^*) < 0\},$$

$$H = \{x \in R^2 / (f_0'(x^*), x - x^*) < 0\}:$$

K բազմության մեջ մտնում են հարթության այն բոլոր x կետերը, որ x^* կետից $h = x - x^*$ վեկտորի ուղղությամբ բավականաչափ փոքր շարժվելիս մնում ենք թույլատրելի լուծումների բազմության մեջ: Իսկ H-ի էլեմենտները այն վեկտորներն են, որ $h = x - x^*$-ի ուղղությամբ շարժվելիս



Գծ. 4.11: Կուն- Տակկերի թեորեմի դիֆերենցիալ տեսքը

f_0 ֆունկցիայի արժեքները փոքրանում են: Այսպիսով, H և K բազմությունները ընդհանուր էլեմենտ չունեն, այսինքն՝

$$K \bigcap H = \emptyset \text{ (տես գծ.4.11)}:$$

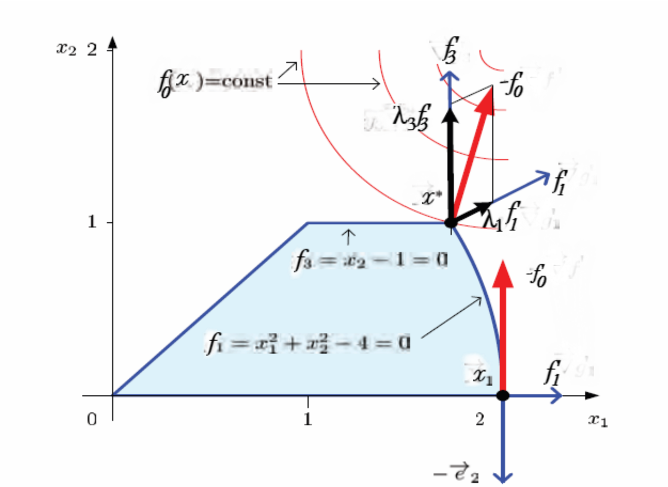
Այժմ համաձայն Թեորեմ 5.5[24]-ի (տես գլուխ 3, էջ117) գոյություն ունեն այնպիսի λ_1 և λ_2 ոչ բասական թվեր, որ

$$f_0'(x^*) + \lambda_1 f_1'(x^*) + \lambda_2 f_2'(x^*) = 0:$$

Սա նշանակում է, որ մինիմումի կետում նպատակային ֆունկ_ցիայի և սահմանափակումներում մասնակցող ֆունկցիաների գրադիենտները պետք է լինեն ոչ բացա– սական գործակիցներով գծորեն կախված, որը Կուն-Տակկերի թեորեմի դիֆերենցիալ տեսքն է: Կամ որ նույն է՝ նպատակային ֆունկցիայի հակագրադիենտը պետք է ընկած լինի ակտիվ սահմանափակումների արտաքին նորմալների վրա ձգված կոնի մեջ:

Այժմ դիտարկենք կոնկրետ մի օրինակ:

Օրինակ: Դիտարկենք ուռուցիկ ծրագրավորման հետևյալ



Գծ. 4.12: Քառակուսային ֆունկցիայի մինիմիզացիան անհավասարությունների փիպի հինգ սահմանափակումների դեպքում

խնդիրը.

$$f_0(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 2)^2 \to min,$$

$$x_1^2 + x_2^2 \leq 4,$$

$$x_2 \leq x_1,$$

$$x_2 \leq 1,\ x_1 \geq 0,\ x_2 \geq 0:$$

Ինչպես երևում է գծ.4.12-ից թույլատրելի վեկտորների բազմությունը կորագիծ սեղան է, իսկ մինիմումի կետը գտնվում է վերևի աջ անկյունում: Այդ կետի կոորդինատները բավարարում են հավասարումների հետևյալ համակարգին.

$$\begin{cases} x_1^2 + x_2^2 = 4, \\ x_2 = 1\ : \end{cases}$$

Այստեղից $x_1^* = \sqrt{3}, \;\; x_2^* = 1$: Կուն- Տակկերի թեորեմի պայմանները գրելու համար սահմանափակումները բերենք կանոնական տեսքի.

$$f_1(x_1, x_2) = x_1^2 + x_2^2 - 4 \leq 0,$$

$$f_2(x_1, x_2) = -x_1 + x_2 \leq 0,$$

$$f_3(x_1, x_2) = x_2 - 1 \leq 0, f_4(x_1, x_2) = -x_1 \leq 0,$$

$$f_5(x_1, x_2) = -x_2 \leq 0:$$

Գծ.4.12-ից երևում է, որ ակտիվ են առաջին և երրորդ սահմանափակումները: Հաշվենք նպատակային ֆունկցիայի և սահմանափակումների գրադիենտները մինիմումի կետում: Ունենք

$$-f_0'(x^*) = (4 - 2\sqrt{3}, 2), \;\; f_1'(x^*) = (2\sqrt{3}, 2), \;\; f_3'(x^*) = (0, 1):$$

Վերլուծելով $-f'(x^*)$ հակագրադիենտը $f_1'(x^*)$, $f_3'(x^*)$ վեկտորով՝ կստանանք

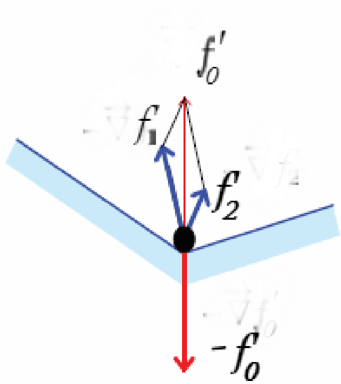
$$-f'(x^*) = \lambda_1 f'(x^*) + \lambda_3 f'(x^*) :$$

Այստեղ

$$\lambda_1 = \frac{2\sqrt{3-2}}{3} > 0, \;\; \lambda_3 = 12 - 4\sqrt{36} > 0:$$

Այսպիսով, $\lambda_1, \;\; \lambda_2$ թվերը դրական են, հետևաբար x^* կետը իրականում խնդրի լուծումն է: Իսկ օրինակ $x^{**} = (2, 0)$ կետում, որը խնդրի լուծումը չէ ակտիվ են առաջին և չորրորդ սահմանափակումները և տեղի ունի

$$-f_0'(x^{**}) = 0 \cdot f_1'(x^{**}) - 4f_4'(x^{**})$$



Գծ. 4.13: Կուն-Տակկերի թեորեմի ֆիզիկական մեկնաբանությունը

հավասարությունը: Այս վերլուծության երկրորդ գործակիցը ստացվեց բացասական: Այդպես էլ պետք է լիներ, քանի որ x^{**}-ը խնդրի լուծումը չէ:

Կուն-Տակկերի թեորեմը ունի նաև շատ պարզ ֆիզիկական մեկնաբանություն: Ենթադրենք ունենք գնդիկ, որը ծանրության $-f_0'$ ուժի ազդեցության տակ շարժվում է մի մակերևույթով, որը բաղկացած է $f_1(x) = 0$, $f_2(x) = 0$ ողորկ մակերևույթներից (տես գծ.4.13):

Ըստ Նյուտոնի երրորդ օրենքի՝ յուրաքանչյուր պահի գնդիկի վրա ազդում է ողորկ այն մակերևույթի հակազդեցության ուժը, որին նա այդ պահին շփվում է (ակտիվ սահմանափակում): Այդ ուժերը ուղղահայաց են մակերևույթներին և հետևաբար հակուղղված են համապատասխանաբար f_1', f_2' գրադիենտներին: Գնդիկը կանգնում է ստացիոնար x^* կետում, որտեղ նրա վրա ազդող բոլոր ուժերի համազորը հավասար է զրոյի.

$$-f_0'(x^*) - \lambda_1 f_1'(x^*) - \lambda_2 f_2'(x^*) = 0,$$

ընդ որում՝ այս վերլուծության λ_1, λ_2 գործակիցները ոչ բացասական են:

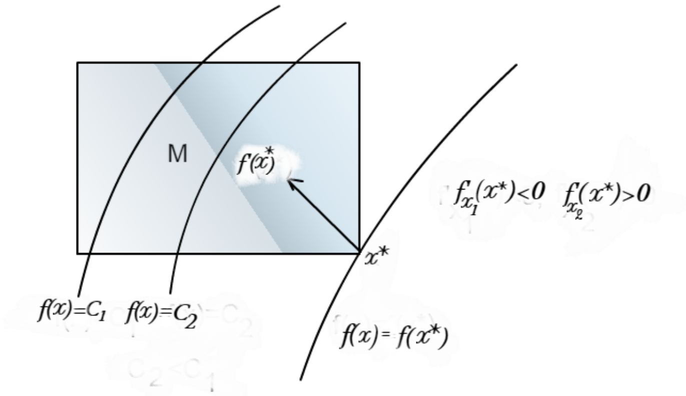
Հետևանք 4.5.1: *Դիցուք*

$$M=\{x\in R^n/a_j\le x_j\le b_j,\quad j\in[1:n]\},$$

որտեղ $-\infty\le a_j<b_j\le+\infty,\ \ j\in[1:n]$*:*
Որպեսզի $x^*\in M$*-ը լինի ստույգ և դիֆերենցելի* $f(x)$ *ֆունկցիայի մինիմումի կետ* M *բազմության վրա, անհրաժեշտ է և բավարար, որ կամայական* $j\in[1:n]$ *համար*

$$f'_{x_j}(x^*)\begin{cases} =0, & \textit{եթե}\ \ a_j<x_j^*<b_j,\\ \ge 0, & \textit{եթե}\ \ x_j^*=a_j\ne-\infty,\\ \le 0, & \textit{եթե}\ \ x_j^*=b_j\ne+\infty: \end{cases}$$

Այս պայմանների երկրաչափական մեկնաբանությունը պրված է գծ.4.14-ում:



Գծ. 4.14: Մինիմումի անհրաժեշտ և բավարար պայմանը, երբ սահմանափակումը ուղղանկյուն է:

Քանի որ Կուն-Տակկերի թեորեմը մինիմումի անհրաժեշտ և բավարար պայման է, ապա որոշ դեպքերում այն թույլ է տալիս բացահայտ տեսքով գտնել անհավասարության տիպի սահմանափակումներով մաթեմատիկական ծրագրավորման խնդիրների լուծումները:

Օրինակ 1: Լուծել հետևյալ խնդիրը.

$$f(x_1,x_2)=2x_1^2+x_1x_2+x_2^2\to min,\ \ -1\le x_1\le 1,\ \ x_2\ge 2:$$

Քանի որ f-ը ուժեղ ուռուցիկ ֆունկցիա է, իսկ սահմա–նափակումներով տրվող բազմությունը փակ է և ուռուցիկ, ապա այս խնդիրը ունի միակ լուծում և այն բավարարում է հետևյալ երկու պայմաններին.

$$f'_{x_1}(x_1,x_2)=4x_1+x_2\begin{cases}=0, & \text{եթե } -1<x_1<1,\\ \geq 0, & \text{եթե } x_1=-1,\\ \leq 0, & \text{եթե } x_1=1,\end{cases}$$

$$f'_{x_2}(x_1,x_2)=x_1+2x_2\begin{cases}=0, & \text{եթե } x_2>2,\\ \geq 0, & \text{եթե } x_2=2:\end{cases}$$

Այժմ այս պայմաններից կարելի է կազմել վեց համակարգեր և լուծել դրանք: Սակայն դժվար չէ համոզվել, որ

$$\begin{cases}4x_1+x_2=0, & \text{եթե } -1<x_1<1,\\ x_1+2x_2\geq 0, & \text{եթե } x_2=2\end{cases}$$

համակարգը համատեղելի է և $(-1/2,2)$ վեկտորը նրա լուծումն է: Հետևաբար այդ վեկտորը նաև մինիմիզացիայի խնդրի միակ լուծումն է:

Օրինակ 2: Լուծել հետևյալ խնդիրը.

$$\begin{array}{l}x_1^2+(x_2-2)^2\to min,\\ x_1^2+x_2^2-1\leq 0,\\ x_1\geq 0,\ x_2\geq 0:\end{array}$$

Քանի որ թույլատրելի արժեքների բազմությունը ունի ներքին կետ, ապա Լագրանժի ֆունկցիայում λ_0 գործակիցը կարելի է վերցնել հավասար մեկի: Կիրառելով Կուն-Տակկերի թեորեմը՝ կստանանք.

$$\begin{cases}L'_{x_1}(x,\lambda)=2x_1+2\lambda_1x_1-\lambda_2=0,\\ L'_{x_2}(x,\lambda)=2(x_2-2)+2\lambda_1x_2-\lambda_3=0,\\ \lambda_1\geq 0,\ \lambda_2\geq 0,\ \lambda_3\geq 0,\\ \lambda_1(x_1^2+x_2^2-1)=0,\ \lambda_2x_2=0,\ \lambda_3x_3=0\\ x_1^2+x_2^2-1\leq 0,\ x_1\geq 0,\ x_2\geq 0:\end{cases}\qquad(4.5.8)$$

Դիտարկենք հետևյալ դեպքերը:

1) $\lambda_1 = 0,\ \ \lambda_2 = 0,\ \ \lambda_3 = 0$: Այդ դեպքում (4.5.8) համակարգից ստանում ենք $x_1 = 0,\ x_2 = 2$, որը չի բավարարում $x_1^2 + x_2^2 - 1 \leq 0$ անհավասարությանը:

2) $\lambda_1 \neq 0,\ \ \lambda_2 = 0,\ \ \lambda_3 = 0$: Այս դեպքում (4.5.8) համակարգից կստանանք

$$\begin{cases} x_1^2 + x_2^2 - 1 = 0, \\ 2x_1(1 + \lambda_1) = 0, \\ 2(x_2 - 2) + 2\lambda_1 x_2 = 0: \end{cases}$$

Եթե $\lambda_1 = -1$, ապա երրորդ հավասարումը տեղի չունի: Եթե $\lambda_1 \neq -1$, ապա կստանանք

$$x_1 = 0,\ x_2 = 1,\ \lambda_1 = 1,\ \lambda_2 = 0:$$

Քանի որ մինիմիզացվող ֆունկցիան ուժեղ ուռուցիկ է, իսկ թույլատրելի վեկտորների բազմությունը փակ է և ուռուցիկ, ապա $(0, 1)$ կետը խնդրի միակ լուծումն է:

4.6 Մաթեմատիկական ծրագրավորման խնդիրների լուծման գրաֆիկական եղանակը

Գրաֆիկական մեթոդով (գծապատկերով) գտնում ենք երկու փոփոխականի $z = f(x_1, x_2)$ ֆունկցիայի մինիմումի (մաքսիմումի) կետերը

$$M = \{x \in R^2 / g_i(x_1, x_2) \leq \alpha_1, i = 1, 2, \ldots k\}$$

բազմության վրա: Կատարում ենք հետևյալ գործողությունները:

- Հարթության վրա կառուցում ենք թույլատրելի արժեքների M բազմությունը:
- Կառուցում ենք նաև f ֆունկցիայի մակարդակի $V_C = \{x \in R^2 / f(x_1, x_2) = C\}$ բազմությունները C պարամետրերի տարբեր արժեքների դեպքում:
- Որոշում ենք մակարդակի գծերի նվազման (աճման) ուղղությունները:
- Գտնում ենք C պարամետրի այն ամենափոքր (ամենամեծ) արժեքը, որի դեպքում

$$M^* \equiv M \bigcap V_C$$

  բազմությունը դատարկ չէ: Այն f ֆունկցիայի մինիմումի (մաքսիմումի) կետերի բազմությունն է:

Մաթեմատիկական ծրագրավորման խնդիրների լուծման գրաֆիկական եղանակը մեկնաբանենք թվային օրինակների միջոցով:

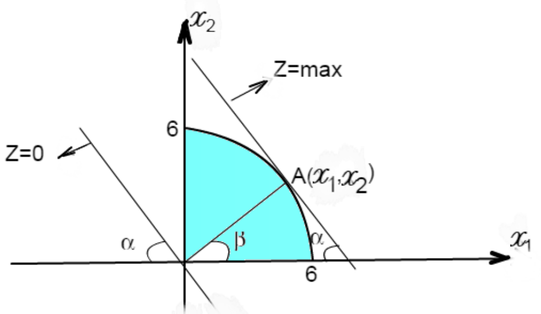
Օրինակ 1: Լուծել հետևյալ խնդիրը.

$$z = f(x_1, x_2) = 2x_1 + x_2 \to \max,$$

$$x_1^2 + x_2^2 \leq 36,$$

$$x_1,\ x_2 \geq 0:$$

Կառուցենք թույլատրելի արժեքների տիրույթը, այսինքն՝ սահմանափակումներով տրվող բազմությունը: Դա 0 կենտրոնով և 6 շառավիղով շրջանի այն մասն է, որը գտնվում է առաջին քառորդում (տես գծ. 4.15): Հաշվի առնելով, որ f-ի մակարդակի բազմությունները $2x_1 + x_2 = C$ գծեր են, ապա պետք է գտնել այն ամենամեծ C թիվը, որի դեպքում այդ գիծը հատում է թույլատրելի արժեքների տիրույթը: Գծ.4.15-ից երևում է, որ այդ դեպքում f ֆունկցիայի մակարդակի բազմությունը և թույլատրելի տիրույթը հատվում են A կետում: Գտնենք այդ կետի կոորդինատները: Ունենք



Գծ. 4.15: Խնդրի գրաֆիկական լուծումը

$$| \, tg\alpha \, |= 2, \; tg\beta = ctg\alpha = \frac{1}{2}:$$

Հետևաբար $A(x_1, x_2)$ կետի կոորդինատները պետք է բավարարեն հավասարումների հետևյալ համակարգին.

$$\begin{cases} x_1^2 + x_2^2 = 36, \\ \frac{x_1}{x_2} = 2 : \end{cases}$$

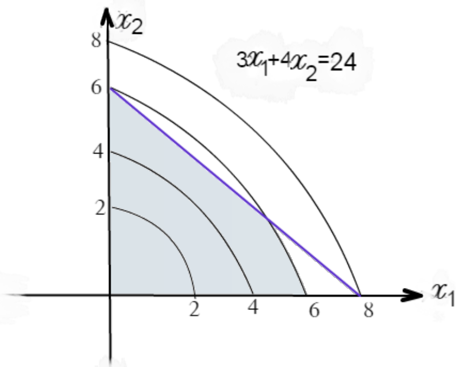
Այստեղից

$$x_1^{max} = \frac{12}{\sqrt{5}}, \; x_2^{max} = \frac{6}{\sqrt{5}}:$$

Օրինակ 2: Լուծել խնդիրը.

$$\begin{cases} z = \sqrt{x_1^2 + x_2^2} \to max, \\ 3x_1 + 4x_2 - 24 \le 0, \\ x_1 \ge 0, x_2 \ge 0 : \end{cases}$$

Թույլատրելի արժեքների տիրույթը 6 և 8 կողերով ուղղանկյուն եռանկյուն է (տես գծ. 4.16): Հաշվի առնելով, որ z ֆունկցիայի մակարդակի գծերը 0 կենտրոնով $\sqrt{c}$ $(c \ge 0)$ շառավղով շրջանագծեր են, ապա պետք է գտնել այն ամենամեծ շառավիղը, որին համապատասխան շրջանագիծը հատում է նշված եռանկյունը: Գծ.4.16-ից երևում է, որ դա 8 շառավիղով շրջանագիծն է: Այստեղից կստանանք $x_1^{max} = 8, x_2^{max} = 0$:



Գծ. 4.16: Խնդրի գրաֆիկական լուծումը

Այժմ բերենք տնտեսագիտական մի խնդիր, որի մաթեմատիկական ձևակերպումը մաթեմատիկական ծրագրավորման կոտորակագծային խնդիր է, և այն լուծենք գրաֆիկական եղանակով:

Օրինակ 3: Դիցուք գործարանը A և B փիփի արտադրանքների արտադրությունը կազմակերպում է երեք փիպի սարքավորումների օգնությամբ: Աղյուսակում (տես գծ.4.17) ներկայացված է A և B տեսակի միավոր արտադրանքների արտադրության ժամանակները և ծախսերը: Ենթադրենք առաջին և երրորդ փիպի սարքավորումները կարելի է օգտագործել համապատասխանաբար ոչ ավելի քան 26 և 39 ժամ: Իսկ երկրորդ սարքավորումը՝ ոչ պակաս 4 ժամ: Պահանջվում է գտնել, թե որքան պետք է արտադրել յուրաքանչյուր փիպի ապրանքից, որպեսզի միավոր արտադրանքի միջին ինքնարժեքը լինի փոքրագույնը:

Կառուցենք այս խնդրի մաթեմատիկական մոդելը: x_1-ով նշանակենք առաջին փիպի արտադրված արտադրանքի քանակը, իսկ x_2-ով՝ երկրորդ փիպի արտադրանքի քանակը: Արտադրության ընդհանուր ֆինանսական ծախսը կլինի՝ $2x_1 + 3x_2$, իսկ միավոր արտադրանքի միջին ինքնարժեքը կլինի

$$\frac{2x_1 + 3x_2}{x_1 + x_2}:$$

Հետևաբար, պետք է լուծել մինիմումի հետևյալ խնդիրը.

$$z = \frac{2x_1 + 3x_2}{x_1 + x_2} \to min,$$

$$\begin{cases} 2x_1 + 8x_2 \leq 26, \\ x_1 + x_2 \geq 4, \\ 12x_1 + 3x_2 \leq 39, \\ x_1,\ x_2 \geq 0 : \end{cases}$$

Գրաֆիկորեն ներկայացնենք խնդրի թույլատրելի արժեքների փիրույթը: Այն պատկերված է ստորև (տես գծ.4.18): Այն ABC եռանկյունն է: Նպատակային ֆունկցիայի բանաձևից ունենք

$$x_2 = \frac{z-2}{3-z} x_1 = k x_1 :$$

Ուղղի k անկյունային գործակիցը հավասար է

$$k = \frac{z-2}{3-z}:$$

Այստեղից

$$\frac{dk}{dz} = \frac{1}{(3-z)^2} > 0,$$

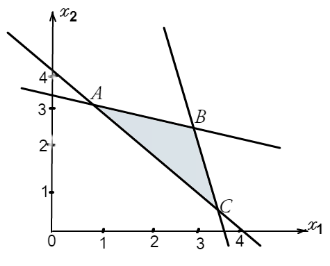
ինչը նշանակում է, որ k ֆունկցիան աճում է և $x_2 = kx_1$ ուղիղը այդ աճման պրոցեսում պտտվում է ժամ սլաքի հակառակ ուղղությամբ: Հետևաբար նպատակային ֆունկցիան իր փոքրագույն արժեքն ընդունում է C կետում: Այդ կետի կոորդինատները գտնելու համար լուծենք հետևյալ համակարգը.

$$\begin{cases} 12x_1 + 3x_2 = 39, \\ x_1 + x_2 = 4: \end{cases}$$

Այստեղից կստանանք՝ $x_1 = 3$, $x_2 = 1$: Հետևաբար օպտիմալ պլանի դեպքում գործարանը պետք է արտադրի 3 միավոր արտադրանք A տիպի ապրանքից և մեկ միավոր՝ B տիպի ապրանքից: Իսկ այդ դեպքում միավոր արտադրանքի մինիմալ ինքնարժեքը հավասար կլինի 2.25 դրամական միավորի:

Սարքավորման տիպը	Միավոր արտադրանքի վրա ծախսված ժամանակը (ժամ)	
	A	B
I	2	8
II	1	1
III	12	3
Դրամական ծախսը միավոր ապրանքի արտադրության վրա (դրամական միավոր)	2	3

Գծ. 4.17: Արտադրության տեխնոլոգիան



Գծ. 4.18: Թույլատրելի արժեքների տիրույթը

4.7 Կուն-Տակկերի թեորեմի կիրառությունը գրադիենտի պրոյեկտման մեթոդում

Այժմ քննարկենք գրադիենտի պրոյեկտման մեթոդը մասնավոր դեպքերում, երբ բազմությունը տրվում է հավասարության կամ անհավասարության փիպի գծային սահմանափակումների փեսքերով:

Քննարկենք գրադիենտի պրոյեկտման մեթոդը, երբ M-ը **աֆինական** բազմություն է, այսինքն՝

$$M = \{x \in R^n / Ax = b\}:$$

Նշանակենք $H = \{x \in R^n / Ax = 0\}$:

Տեղի ունի հետևյալ պնդումը:

Լեմմ 4.7.1 *Դիցուք $A(m \times n)$ մատրիցի տողերը գծորեն անկախ են: Այդ դեպքում y կետի պրոյեկցիան H բազմության վրա որոշվում է հետևյալ բանաձևով.*

$$\Pi_H(y) \equiv x = (I - P)y,$$

$$P = A^T(AA^T)^{-1}A:$$

► Նախ ցույց տանք, որ (AA^T) մատրիցը հակադարձելի է: Ենթադրենք հակառակը: Այդ դեպքում գոյություն կունենա այնպիսի ոչ զրոյական z վեկտոր, որ $(AA^T)z = 0$: Այստեղից

$$z^T(AA^T)z = (zA^T)^T A^T z = (A^T z, A^T z) = \|A^T x\|^2 = 0,$$

այսինքն՝ $A^T z = 0$: Բայց քանի որ A մատրիցի տողերը գծորեն անկախ են, ապա այստեղից կունենանք $z = 0$, ինչը հակասություն է: P մատրիցն ունի հետևյալ հատկությունները.

$$P^2 = P,\ P(I - P) = 0, P^T = P:$$

Այստեղից հետևում է, որ կամայական y վեկտորի համար տեղի ունեն

$$y = Py + (I - P)y,$$

$$(Py, (I - P)y) = (y, P(I - P)y = 0$$

հավասարությունները: Այսպիսով, յուրաքանչյուր y վեկտոր հավասար է երկու օրթոգոնալ վեկտորների գումարի, որոնցից մեկը y-ի պրոյեկցիան է H-ի վրա, մյուսը պատկանում է նրա օրթոգոնալ լրացմանը: Դիցուք այժմ y-ը կամայական վեկտոր է, իսկ $x \in H^{\perp}$: Այդ դեպքում

$$\|y - x\|^2 = \|Py - x + (I - P)y\|^2 = \|Py - x\|^2 +$$

$$+2(Py - x, (I - P)y) + \|(I - P)y\|^2:$$

Քանի որ $x \in H^{\perp}$, $(I - P)y \in H$, ապա

$$(Py, (I - P)y) = 0, \ (x, (I - P)y) = 0:$$

Այնպես որ

$$\|y - x\|^2 = \|Py - x\|^2 + \|(I - P)y)\|^2:$$

Այստեղից հետևում է, որ y կետի հեռավորությունը $H^{\perp}$ բազմությունից կլինի փոքրագույն, եթե $x = Py$:■

Օրինակ: Գտնել $y = (3, 0)$ կետի պրոյեկցիան $M = \{x \in R^2 / x_1 + 2x_2 = 4\}$ գծի վրա: Այստեղ $A = (1, 2)$, իսկ պրոյեկտման մատրիցան հետևյալն է.

$$I - A^T(AA^T)^{-1}A = \begin{pmatrix} 0.8, & -0.4 \\ -0.4 & 0.2 \end{pmatrix}:$$

Այսպեղից՝ $\Pi_H(y) = (I - P)y = (2.4, -1.2)$: Հետևաբար $h = \Pi_M(y) = (0, 2) + (2.4, -1.2) = (2.4, 0.8)$:

Այժմ կազմենք խնդրի Լագրանժի ֆունկցիան.

$$L(x, \lambda) = f(x) + \lambda^T(Ax - b):$$

Այս դեպքում Կուն-Տակկերի թեորեմի դիֆերենցելի պայմանները կունենան հետևյալ տեսքը.

$$L'_x(x, \lambda) = f'(x) + A^T\lambda = 0, \ Ax = 0:$$

Դիցուք գտնվում ենք x^k կետում: Այդ կետից որպես շարժման ուղղություն վերցվում է $-f'(x^k)$ վեկտորի պրոյեկցիան H բազմության վրա, այսինքն՝ $h^k = -(I - P)f'(x^k)$ վեկտորը: Հաջորդ կետը որոշում ենք հետևյալ ռեկուրենտ առնչությամբ.

$$x^{k+1} = x^k + \alpha_k h^k:$$

Այս ռեկուրենտ առնչությունը օժտված է հետևյալ կարևոր հատկություններով.

ա) կամայական $\alpha_k > 0$ թվերի համար x^{k+1}-ը պատկանում է թույլատրելի կետերի բազմությանը: Իրոք,

$$Ax^{k+1} = A(x^k + \alpha_k h^k) = Ax^k + \alpha_k Ah^k =$$

$$= b - \alpha_k A(I - A^T(AA^T)^{-1}A)f'(x^k) =$$

$$= b - \alpha_k(A - A)f'(x^k) = b:$$

բ) Եթե $(I-P)f'(x^k)=0$, ապա x^k-ն խնդրի լուծումն է: Իրոք,

$$(I-P)f'(x^k)=(I-A^T(AA^T)^{-1}A)f'(x^k)=$$

$$=f'(x^k)+A^T(-(AA^T)^{-1}Af'(x^k))=0:$$

Նշանակելով $\lambda^k=-(AA^T)^{-1}Af'(x^k)$, ստանում ենք, որ x^k կետը բավարարում է Կուն-Տակկերի թեորեմի պայմաններին.

$$f'(x^k)+A^T\lambda^k=0,\ Ax^k=b:$$

Հետևաբար x^k-ն խնդրի լուծումն է: α_k քայլի երկարությունը ընտրվում է ամենաարագ վայրէջքի մեթոդով:

գ) Եթե $h^k\neq 0$, ապա այն ֆունկցիայի նվազման ուղղությունն է: Դիտարկենք մեկ փոփոխականի հետևյալ ֆունկցիան.

$$\varphi(\alpha)\equiv f(x^k+\alpha h^k):$$

Հաշվենք նրա ածանցյալը 0 կետում: Ըստ բարդ ֆունկցիայի ածանցման կանոնի՝ հաշվի առնելով նաև P օպերատորի վերը նշված հատկությունները՝ կստանանք

$$\varphi'(0)=-(f'(x^k),h^k)=-(f'(x^k),(I-P)f'(x^k))=$$

$$-(f'(x^k),(I-P)^2f'(x^k))=$$

$$=(f'(x^k),(I-P)^T(I-P)f'(x^k))=$$

$$=-((I-P)f'(x^k),(I-P)f'(x^k))<0:$$

Այսպեղից՝ քանի որ $\varphi'(0) < 0$, ապա բավականաչափ փոքր դրական α թվերի համար $\varphi(\alpha) < \varphi(0)$:

Ուրեմն h^k-ն f-ի նվազման ուղղությունն է:

Այժմ քննարկենք գրադիենտի պրոյեկցման մեթոդը, երբ M-ը տրված է գծային անհավասարությունների միջոցով:

Դիտարկենք ուռուցիկ ծրագրավորման հետևյալ խնդիրը.

$$f(x) \to min, Ax \leq b:$$

Օպտիմալության x կետում տեղի ունեն Կուն-Տակկերի թեորեմի պայմանները.

1) $L'_x(x, \lambda) = f'(x) + A^T\lambda = 0,$

2) $Ax - b \leq 0,$

3) $\lambda_i(Ax - b)_i = 0,\ i = 1, 2, .., m,$

4) $\lambda \geq 0 :$

Նկարագրենք գրադիենտի պրոյեկցման մեթոդը այս խնդրի համար:

- **Առաջին քայլում** ընթացիկ x^k կետում ստուգում ենք Կուն-Տակկերի օպտիմալության պայմանները: Դրա համար նախապես որոշում ենք ակտիվ սահմանափակումները: Տեղադրում ենք x^k կետը անհավասարությունների մեջ: Ենթադրենք առաջին q հատը դառնում են հավասարություններ, իսկ մնացած անհավասարությունները խիստ են.

$$a_i^T x^k = b_i,\ i = 1, 2, \ldots, q,$$

$$a_i^T x^k < b_i,\ i = q+1, \ldots, m:$$

Այստեղ $a_i^T,\ i = 1, 2, \ldots, m$ վեկտորները $A(n \times m)$ մատրիցի տողերն են: Այսպիսով, ակտիվ են առաջին q սահմանափակումները:

Եթե ակտիվ սահմանափակումներ կան ($q \geq 1$), ապա նրանց հիմքի վրա կառուցում ենք պրոյեկտման հետևյալ օպերատորը.

$$\hat{P}_q = I - A_q^T (A_q A^T)^{-1} A_q,$$

որտեղ A_q-ն $a_i^T,\ \ i = 1, 2, \ldots, q$ տողերից բաղկացած մատրիցն է: Համաձայն Կուն-Տակկերի թեորեմի՝ x^k կետը կլինի օպտիմալ այն և միայն այն դեպքում, երբ տեղի ունենան հետևյալ երկու պայմանները.

ա) $\hat{P}_q f'(x^k) = 0$:

բ) Ակտիվ սահմանափակումների համապատասխանող Լագրանժի գործակիցները ոչ բացասական են՝ $\lambda = -(A_q A_q^T)^{-1} A_q f'(x^k) \geq 0$: Իսկ եթե ակտիվ սահմանափակում չկա, ապա օպտիմալության պայմանը ունի պարզ տեսք՝ $f'(x^k) = 0$:

- **Երկրորդ քայլում,** եթե օպտիմալության պայմանը տեղի չունի, ապա պետք է շարժվել այնպիսի ուղղությամբ, որ ֆունկցիայի արժեքը փոքրանա, միաժամանակ մնալով բազմության մեջ: Այստեղ հնարավոր են հետևյալ տարբերակները:

  Տարբերակ առաջին: Եթե ակտիվ սահմանափակումները բացակայում են, ապա որպես շարժման ուղղություն վերցվում է հակագրադիենտի ուղղությունը. $h^k = -f'(x^k)$:

Տարբերակ երկրորդ: $\hat{P}_q f'(x^k) \neq 0$: Այս դեպքը համապատասխանում է այն իրավիճակին, երբ $1 \geq q < n$:

Այս դեպքում x^k կետը գտնվում է եզրի վրա, որը հանդիսանում է աֆինական բազմություն և այդ կետից՝ որպես ֆունկցիայի նվազման ուղղություն, ընտրվում է $h^k = \hat{P}_q f'(x^k)$ վեկտորը:

Տարբերակ երրորդ: $\hat{P}_q f'(x^k) = 0$ և Լագրանժի $\lambda_1, \ldots, \lambda_q$ գործակիցներից մեկը բացասական է: Այս դեպքը տեղի ունի այն ժամանակ, երբ $q = n$: Հետևաբար աֆինական բազմությունը, որի վրա պետք է կատարել պրոյեկտում մի կետից է բաղկացած: Այս դեպքում ակտիվ սահմանափակումների ցուցակից հեռացնում ենք այն սահմանափակումը, որին համապատասխանող Լագրանժի գործակիցը բացասական է: Հեռացնելուց հետո ստացվում է A_{q-1} մատրիցը և կառուցում ենք պրոյեկտման նոր մատրից.

$$\hat{P}_{q-1} = I - A_{q-1}^T (A_{q-1} A_{q-1}^T)^{-1} A_{q-1}:$$

Այնուհետև որպես շարժման ուղղություն ընտրում ենք $h^k = -\hat{P}_{q-1} f'(x^k)$ վեկտորը:

- **Երրորդ քայլում** որոշում ենք շարժման ուղղությամբ α_k քայլի երկարությունը: Նախ որոշում ենք, թե այդ քայլը վերևից ինչով է սահմանափակ: Դա այն T պահն է, երբ $x^k + \alpha h^k$ ճառագայթը դուրս է գալիս թույլատրելի արժեքների տիրույթից: Գծային սահմանափակումների դեպքում այդ թիվը հեշտությամբ որոշվում է: Եթե i-րդ սահմանափակումը ակտիվ չէ, ապա այն T_i պահը, երբ ճառագայթը հատում է $(a_i, x) = b_i$ հիպերհարթությունը որոշվում է հետևյալ հավասարումից.

$$(a_i, x^k + \alpha h^k) = b_i:$$

Որտեղից,

$$T_i = \frac{b_i - (a_i, x^k)}{(a_i, h^k)}:$$

Իսկ այն T պահը երբ ճառագայթը առաջին անգամ հատում է պասիվ սահմանափակումները որոշվում է հետևյալ բանաձևով.

$$T = \min_{i \in [q+1:m]} max(0, T_i):$$

Վերջապես α_k քայլի երկարությունը կարելի է որոշել արագ վայրէջքի մեթոդով.

$$\alpha_k = arg \min_{0 \le \alpha \le T} f(x^k + \alpha h^k):$$

Այժմ դիտարկենք օրինակ, որի միջոցով լուսաբանենք վերևում քննարկված ալգորիթմը:

Օրինակ:

$$F(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2 \to min,$$

$$x_1 + x_2 \le 2,$$

$$x_2 \le 1,$$

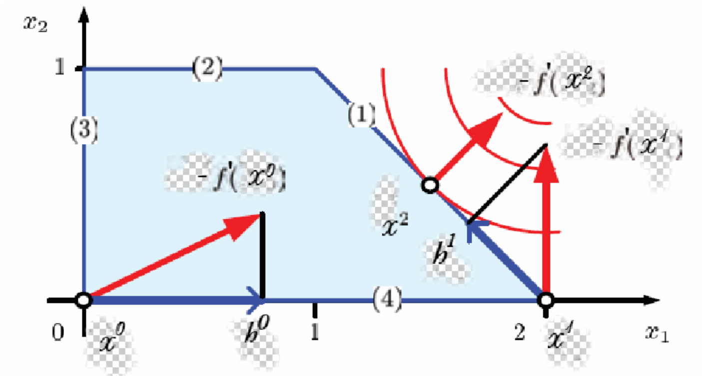
$$-x_1 \quad\quad \le 0,$$

$$-x_2 \leq 0:$$

Այս օրինակում

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}:$$

Խնդրի երկրաչափական լուծումը բերված է գծանկարում:



Գծ. 4.19: Գրադիենտի պրոյեկցման մեթոդը գծային սահմանափակումների դեպքում: Սահմանափակումները նշված են համարներով փակագծերի մեջ:

Որպես սկզբնական մոտավորություն վերցնում ենք $x^0 = (0,0)$ կետը:

- **Առաջին իտերացիայում** որոշում ենք x^0 կետում ակտիվ սահմանափակումները: Տեղադրելով անհավասարությունների մեջ՝ պեսնում ենք, որ ակտիվ են երրորդ և չորրորդ սահմա–նափակումները: Հետևաբար

$$A_q = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}:$$

Քանի որ A_q մատրիցը քառակուսային է, ապա պրոյեկտման $\hat{P}_q = I - A_q^T(A_qA^T)^{-1}A_q$ օպերատորը զրոյական է: Որոշենք Լագրանժի գործակիցները.

$$\lambda = -(A_qA_q^T)^{-1}A_qf'(x^0) = -(A^T)^{-1}A_q^{-1}A_qf'(x^0) =$$

$$-(A^T)^{-1}f'(x^0) = (-4,-2):$$

Երկու գործակիցն էլ բացասական են, ուստի x^0 կետը օպտիմալ չէ: Հակագրադիենտը $-f'(x^0)$ վեկտորն է, որը ուղղված է փիրույթի ներսում: Ակտիվ սահմանափակումների ցուցակից հեռացնենք օրինակ երրորդը: Ստանում ենք մի փոքրից բաղկացած $A_{q-1} = (0,-1)$ մատրիցը: Հաշվենք պրոյեկտման օպերատորը.

$$\hat{P}_{q-1} = I - A_{q-1}^T(A_{q-1}A_{q-1}^T)^{-1}A_{q-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}:$$

Այսպիսով, շարժման ուղղությունը կլինի

$$h^0 = -\hat{P}_{q-1}f'(x^0) = (4,0)$$

վեկտորը: Քայլի մաքսիմալ T երկարությունը հաշվելով՝ ստանում ենք $T = 0.5$: Իսկ քայլի α_0 երկարությունը հաշվում ենք հետևյալ բանաձևով.

$$\alpha_0 = arg \min_{0\le\alpha\le T} f(x^0 + \alpha h^0) = 0.5:$$

- **Երկրորդ իտերացիայում** մեկ քայլից հետո տեղափոխվում ենք $x^1 = x^0 + \alpha_0 h^0 = (2,0)$ կետը: Այս նոր կետում ակտիվ են առաջին և չորրորդ սահմանափակումները, իսկ նրանց համապատասխանող A_q մատրիցը հետևյալն է.

$$A_q = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}:$$

Քանի որ մատրիցը քառակուսային է, ապա պրոյեկտման օպերատորը զրոյական է: Որոշում ենք ակտիվ սահմանափակումներին համապատասխանող Լագրանժի գործակիցները.

$$\lambda = -(A_q^T)f'(x^1) = (0, -2):$$

Քանի որ երկրորդ գործակիցը բացասական է, ուրեմն x^1 կետը օպտիմալ չէ: Ակտիվ սահմանափակումների ցուցակից արտաքսում ենք չորրորդ սահմանափակումը: Ստանում ենք մի տողանի $A_{q-1} = (1, 1)$ մատրիցը, իսկ համապատասխան պրոյեկտման օպերատորը կլինի

$$\hat{P}_{q-1} = I - A_{q-1}^T(A_{q-1}A_{q-1}^T)^{-1}A_{q-1} =$$
$$= \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}:$$

Այնուհետև պրոյեկտում ենք հակագրադիենտը առաջին սահմանափակման վրա և ստանում ենք շարժման ուղղությունը.

$$h^1 = -\hat{P}_{q-1}f'(x^1) = (-1, 1):$$

Նույն կերպ ինչպես առաջին իտերացիայում ստանում ենք քայլի երկարությունը՝ $\alpha_1 =$

0.5: Կատարելով համապատասխան քայլ՝ տեղափոխվում ենք

$$x^2 = x^1 + \alpha_1 h^1 = (1.5, 0.5)$$

կետը:

- **Երրորդ իտերացիայում** ստուգում ենք, որ x^2 կետում ակտիվ է միայն առաջին սահմանա–փակումը: Նրան համապատասխան մատրիցը հետևյալն է. $A_q = (1, 1)$, իսկ պրոյեկտման օպերատորը կլինի՝

$$\hat{P}_q = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}:$$

Հետևաբար

$$\hat{P}_q f'(x^2) = (0, 0),$$

$$\lambda = -(A_q A_q^T)^{-1} A_q f'(x^2) = 1:$$

Այսպիսով, Կուն Տակկերի թեորեմի պայմանները կատարվում են, ուստի x^2-ը մինիմումի կետ է:

ԽՆԴԻՐՆԵՐ

1. Կուն-Տակկերի թեորեմի օգնությամբ լուծել հետև– յալ խնդիրները.

ա)

$$
\begin{aligned}
&(x_1+2)^2+(x_2+2)^2 \to min,\\
&x_1^2+x_2^2 \le 1,\\
&x_1 \ge 0,\ x_2 \ge 0:
\end{aligned}
$$

բ)

$$
\begin{aligned}
&x_1^2+x_2^2 \to min,\\
&2x_1^2+(x_2-4)^2 \le 1:
\end{aligned}
$$

գ) $4x_1^2 - x_1x_2 + 4x_2^2 \to min,\ 4 \le x_1 \le 8,\ -1 \le x_2 \le$ ≤ 2:

2. Ստուգել, արդյո՞ք $(0,4)$ վեկտորը հետևյալ խնդրի լուծումն է, թե ոչ.

$$
\begin{aligned}
&x_1^2+(x_2-2)^2 \to min,\\
&x_1^2+2(x_2-2)^2 \le 8,\\
&x_1^2+2x_2^2 \le 32:
\end{aligned}
$$

3. Ստուգել, արդյո՞ք $(0,1)$ վեկտորը հետևյալ խնդրի լուծումն է, թե ոչ.

$$
\begin{aligned}
&\exp(x_1-x_2) \to min,\\
&x_1+x_2 \le 1,\\
&x_1 \ge 0,\ x_2 \ge 0:
\end{aligned}
$$

4. Գրաֆիկորեն լուծել մաթեմատիկական ծրագրա– վորման հետևյալ խնդիրը.

$$\begin{aligned}&x_2(x_1 - 5) \to extr,\\&x_1^2 + x_2^2 \le 3:\end{aligned}$$

5. Գրաֆիկորեն լուծել մաթեմատիկական ծրագրա– վորման հետևյալ խնդիրը.

$$\begin{aligned}&(x_1 - 4)^2 + (x_2 - 2)^2 \to min,\\&x_1 + x_2 \le 3,\\&x_1 + 2x_2 \le 4:\end{aligned}$$

6. Օգտագործելով Կուն-Տակկերի թեորեմի պայ– մանները՝ գրադիենտի պրոյեկտման մեթոդով լուծել հետևյալ խնդիրը.

$$\begin{aligned}&(x_1 + 1)^2 + (x_2 - 4)^2 + 1 \to min,\\&2x_1 - x_2 - 2 \le 0,\\&x_1 \ge 0\ x_2 \ge 0:\end{aligned}$$

Գլուխ 5

Լագրանժի անորոշ գործակիցների մեթոդը

Լագրանժի անորոշ գործակիցների մեթոդը օպտի– միզացիայի այն հիմնարար սկզբունքներից է, որի միջոցով պայմանական օպտիմիզացիայի խնդիրները բերվում են ոչ պայմանական օպտիմիզացիայի խնդիրների: Առաջին անգամ այս սկզբունքն առաջարկվել է Լագրանժի կողմից 1797 թվականին և այն հիմնավորվել է հավասարության տիպի սահմանափակումներով ողորկ օպտիմիզացիոն խն– դիրների համար:

Այս գլխում դիտարկվում են մաթեմատիկական ծրա– գրավորման այնպիսի խնդիրներ, որոնցում սահմա– նափակումները տրվում են հավասարություններով և անհավասարություններով: Ենթադրվում է, որ նպատակային ֆունկցիան և սահմանափակումները ներկայացնող ֆունկ– ցիաները ողորկ են: Ցույց է տրվում, որ այդպիսի խնդիրների էքստրեմալները գտնվում են Լագրանժի ֆունկցիայի ստացիոնար կետերի բազմության մեջ:

Վերջին ժամանակներս, փորձ է արվում հիմնավորել Լագրանժի մեթոդը մաթեմատիկական ծրագրավորման ոչ ողորկ խնդիրների համար: Նման ուսումնասիրությունների հետ ընթերցողը կարող է ծանոթանալ [12, 22] աշխատանքներում:

5.1 Լագրանժի անորոշ գործակիցների մեթոդը (հավասարության տիպի սահմանափակումների դեպքը)

Սահմանում 5.1.1: *$h \in R^n$ վեկտորը կոչվում է M բազմության $x^* \in M$ կետում* ***շոշափող***, *եթե գոյություն ունի այնպիսի*

$$\varphi : [-1, 1] \to R^n, \;\; \varphi(\alpha) = o(\alpha)$$

արտապատկերում, որ բավականաչափ փոքր $\alpha > 0$ թվերի համար տեղի ունի հետևյալ պայմանը՝

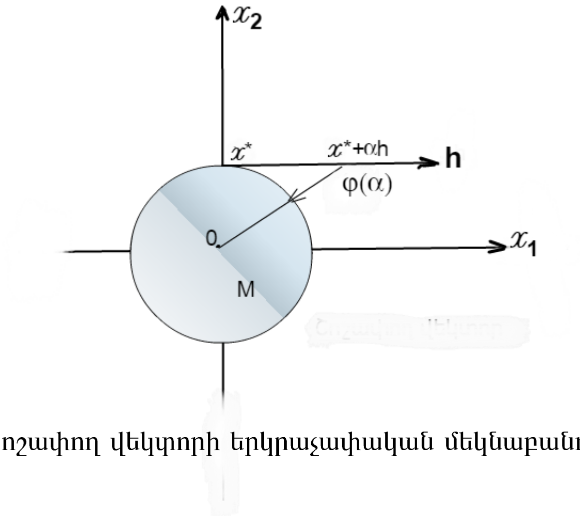
$$x^* + \alpha h + \varphi(\alpha) \in M \; :$$

Օրինակ: Ցույց տանք, որ x առանցքի կամայական h վեկտոր հանդիսանում է շոշափող միավոր $S_1(0)$ շրջանագծին $x^* = (0, 1)$ կետում: Պարզության համար ենթադրենք, որ $\|h\| = 1$: Շրջանագծին տարված շոշափողի և հատողի մասին հայտնի թեորեմի համաձայն՝ ունենք

$$\|\varphi(\alpha)\|(2 + \|\varphi(\alpha)\|) = \alpha^2 \text{ (տես գծ. 5.1) } :$$

Այստեղից կստանանք՝

$$\|\varphi(\alpha)\| = \sqrt{1 + \alpha^2} - 1 = \frac{\alpha^2}{1 + \sqrt{1 + \alpha^2}} = o(\alpha) \; :$$



Գծ. 5.1: Շոշափող վեկտորի երկրաչափական մեկնաբանությունը

$K_M(x^*)$ սիմվոլով նշանակենք M բազմությանը x^* կետում փարված շոշափող վեկտորների բազմությունը:

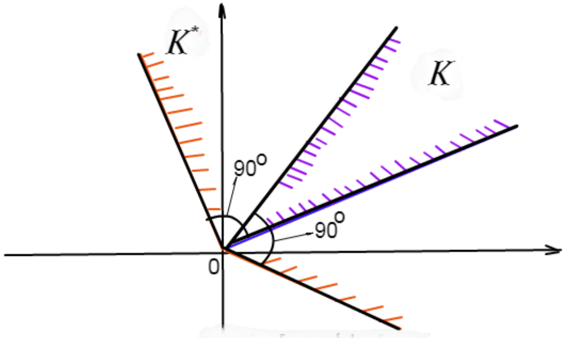
Սահմանում 5.1.2: *Եթե K-ն կոն է, ապա*

$$K^* \equiv \{y \in R^n/(y,x) \geq 0 \ \ \forall x \in K\}$$

բազմությունը կոչվում է K-ի համալուծ կոն:

Այսինքն՝ K^* համալուծ կոնի մեջ մտնում են այն վեկտորները, որոնք K-ի բոլոր էլեմենտների հետ սուր անկյուն են կազմում:

Գծ.5.2-ում բերված է համալուծ կոնի օրինակ:



Գծ. 5.2: Համալուծ կոնի երկրաչափական մեկնաբանությունը

Թեորեմ 5.1.1 *(Լյուստերնիկի թեորեմը շոշափող հարթության մասին)*:*Դիցուք*

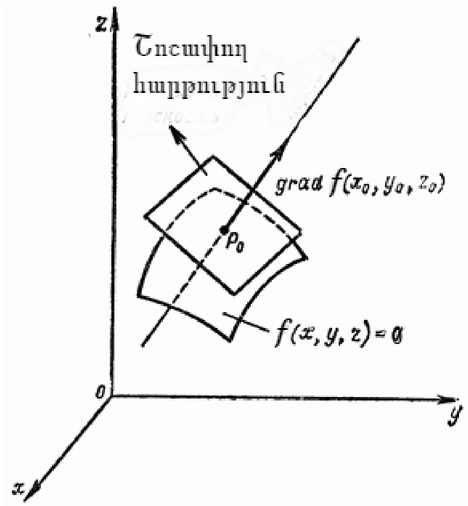
$$M = \{x \in R^n / f_i(x) = 0,\ i \in [1:m]\},$$

որտեղ f_i, $i \in [1:m]$, *ֆունկցիաները անընդհատ դիֆերենցելի են:*
Դիցուք $x^* \in M$ *և ենթադրենք, որ* $f_i'(x^*)$, $i \in [1:m]$ *գրադիենտները գծորեն անկախ են:*
Այդ դեպքում

$$K_M(x^*) = H \equiv \{h \in R^n / \ (f_i'(x^*), h) = 0,\ i \in [1:m]\}:$$

Այսինքն՝ H *ենթապարածությունը շոշափող կոն է* M *բազմության համար* x^* *կետում:*

Գծ.5.3-ում պատկերված է եռաչափ պարածության մեջ $f(x, y, z) = 0$ մակերևույթին $P(x_0, y_0, z_0)$ կետում տարված շոշափող հարթությունը:



Գծ. 5.3: Շոշափող հարթություն

► Ցույց տանք, որ H հարթության կամայական վեկտոր շոշափող վեկտոր է:
Նշանակենք

$$\mathbf{A} = \begin{pmatrix} f'_{1x_1}(x^*) & f'_{1x_2}(x^*)... & f'_{1x_n}(x^*) \\ f'_{2x_1}(x^*) & f'_{2x_2}(x^*)... & f'_{2x_n}(x^*) \\ ... & ... & ... \\ f'_{mx_1}(x^*) & f'_{mx_2}(x^*)... & f'_{mx_n}(x^*) \end{pmatrix}:$$

Կազմենք հավասարումների հետևյալ համակարգը.

$$g_i(\alpha, r) = f_i(x^* + \alpha h + \mathbf{A}^\top r) = 0, \ i \in [1:m], \qquad (5.1.1)$$

որտեղ r-ը m չափանի վեկտոր է: Ցույց տանք, որ այս համակարգը բավարարում է անբացահայտ ֆունկցիայի վերաբերյալ թեորեմի բոլոր պայմաններին:

Իրոք,

$$g_i(0,0)=0,\ \ g'_{i\alpha}(0,0)=(f'_i(x^*),h)=0,\ \ i\in[1:m]\text{:}$$

Ակնհայտ է նաև, որ $\{g'_{ir_j}\}$ կոորդինատներով ֆունկցիոնալ մատրիցը $(0,0)$ կետում հավասար է $\mathbf{B}\equiv\mathbf{A}\mathbf{A}^\top$ մատրիցին, որն ունի հակադարձ:

Հետևաբար գոյություն ունի $r(\alpha)$ արտապատկերում, որոշված զրո կետի ինչ-որ շրջակայքում, այնպիսին, որ այն այդ շրջակայքում բավարարում է (5.1.1) համակարգին և

$$r(0)=0,\ r'(0)=-B^{-1}g'_\alpha(0,0)=0,$$

որտեղ

$$g'_\alpha(0,0)=(g'_{1\alpha}(0,0),...,g'_{m\alpha}(0,0))\text{:}$$

Հետևաբար

$$\lim_{\alpha\to0}\frac{r(\alpha)}{\alpha}=\lim_{\alpha\to0}\frac{r(\alpha)-r(0)}{\alpha}=r'(0)=0\text{:}\qquad(5.1.2)$$

Նշանակելով $\varphi(\alpha)=\mathbf{A}^\top r(\alpha)$, (5.1.1)-(5.1. 2) պայման-ներից՝ կստանանք

$$\varphi(\alpha)=o(\alpha)\text{ և }g_i(x^*+\alpha h+\varphi(\alpha))=0\ \forall i\in[1:m]\text{:}$$

Իսկ սա նշանակում է, որ $h\in K_M(x^*)$:
Այսպիսով, ապացուցվեց, որ $H\subseteq K_M(x^*)$:
Այժմ ցույց տանք, որ $K_M(x^*)\subseteq H$: Դիցուք $h\in$ $\in K_M(x^*)$: Ըստ շոշափող վեկտորի սահմանման գոյություն ունի այնպիսի $\varphi(\alpha)=o(\alpha)$ արտապատկերում, որ բավականաչափ փոքր դրական α թվերի համար

$$f_i(x^*+\alpha h+\varphi(\alpha))=0,\ i\in[1:m]\text{:}$$

Այսպեղից, համաձայն դիֆերենցելիության պայմանի՝ կամայական i-ի դեպքում կունենանք

$$0 = f_i(x^* + \alpha h + \varphi(\alpha)) - f_i(x^*) = \alpha[(f_i'(x^*), h) + \frac{o(\alpha)}{\alpha}]:$$

Ուստի

$$(f_i'(x^*), h) + \frac{o(\alpha)}{\alpha} = 0 \Rightarrow (f_i'(x^*), h) = 0 \Rightarrow h \in H: \blacksquare$$

Հետևանք 5.1.1: *Այս թեորեմից հետևում է, որ f ֆունկցիայի գրադիենտը x^* կետում ուղղահայաց է մակարդակի $V_{f(x^*)}$ բազմությանը:*

Ինչպես տեսաք, այս փաստը կարևոր դեր խաղաց ֆունկցիայի մինիմիզացիայի գրադիենտային մեթոդների մշակման ընթացակարգերում: Դիտարկենք պայմանական օպտիմիզացիայի ընդհանուր խնդիրը.

$$f(x) \to min, \ x \in M: \qquad (5.1.3)$$

Թեորեմ 5.1.2 *(Մինիմումի ընդհանուր անհրաժեշտ պայմանը)*: *Դիցուք f ֆունկցիան դիֆերենցելի է, իսկ $x^* \in M$ կետը (5.1.3) խնդրի լուծում է:*
Այդ դեպքում

$$f'(x^*) \in K_M^*(x^*):$$

► Ենթադրենք

$$f'(x^*) \notin K_M^*(x^*):$$

Այդ դեպքում գոյություն ունի այնպիսի $h \in K_M(x^*)$ վեկտոր, որ $(f'(x^*), h) < 0$: Այստեղից՝ քանի որ f-ը դիֆերենցելի է, ապա բավականաչափ փոքր $\alpha > 0$ թվերի համար կունենանք

$$f(x^* + \alpha h) = f(x^*) + \alpha[(f'(x^*), h) + \frac{o(\alpha)}{\alpha}] < f(x^*):$$

Սա հակասություն է, քանի որ x^*-ն f-ի լոկալ մինիմումի կետ է M բազմության վրա: ■

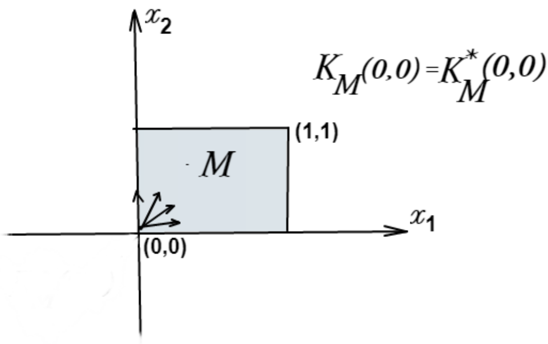
Այս թեորեմը մեկնաբանենք օրինակի միջոցով:

Օրինակ: Դիցուք $f(x_1, x_2) = x_1 + x_2$, իսկ

$$M = \{x \in R^2 / 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\}:$$

Ակնհայտ է, որ $(0, 0)$ կետը f-ի մինիմումի կետն է M-ի վրա: Հեշտ է տեսնել, որ այստեղ $K_M(0, 0)$-ն առաջին քառորդն է, և հետևաբար $K_M^*(0, 0)$ կոնը նույնպես առաջին քառորդն է: Մինիմումի ընդհանուր անհրաժեշտ պայմանը նույնպես տեղի ունի, քանի որ $f'(0, 0) = (1, 1)$, որը առաջին քառորդի կետ է (տես գծ.5.4):

Այսպիսով, էքստրեմումի կոնկրետ խնդրի համար մինիմումի անհրաժեշտ պայման ստանալու համար պետք է սահմանափակումներով տրված բազմության համար կառուցել շոշափող ուղղությունների կոնը և գտնել նրա համալուծը: Ստորև բերվում է համալուծ կոնի կառուցման մի օրինակ, երբ բազմությունը տրված է հավասարության տիպի սահմանափակումներով, իսկ այդ սահմանափակումներում մասնակցող ֆունկցիաները անընդհատ դիֆերենցելի են: Այնուհետև գրելով մինիմումի ընդհանուր անհրաժեշտ պայմանը՝ անմիջապես ստանում ենք գրականության մեջ հայտնի Լագրանժի անորոշ գործակիցների մեթոդը, որն առանցքային դեր ունի էքստրեմումի անհրաժեշտ պայմանների տեսության մեջ:



Գծ. 5.4: Մինիմումի ընդհանուր անհրաժեշտ պայմանի երկրաչափական մեկնաբանությունը

Թեորեմ 5.1.3 *(Համալուծ կոնի կառուցումը): Դիցուք տեղի ունեն **Թեորեմ 5.1.1-ի** պայմանները:*
Այդ դեպքում՝

$$K_M^*(x^*) = H^\perp = A \equiv \{y \in R^n / y = \sum_{i=1}^{m} \lambda_i f_i'(x^*),$$

$$\lambda_i \in R,\ i \in [1:m]\} :$$

► Նախ, քանի որ H-ը ենթապարածություն է և $K_M(x^*) = H$, ապա

$$K_M^*(x^*) = H^\perp :$$

Այժմ ցույց տանք, որ

$$A \subseteq H^\perp:$$

Իրոք, դիցուք

$$\forall h \in H,\ \forall y \in A \Rightarrow y = \sum_{i=1}^{m} \lambda_i f_i'(x^*) \Rightarrow$$

$$\Rightarrow (y, h) = \sum_{i=1}^{m} \lambda_i (f_i'(x^*), h) = 0 \Rightarrow y \in H^{\perp}:$$

Այժմ ենթադրենք, որ գոյություն ունի այնպիսի $b \in H^{\perp}$ վեկտոր, որ $b \notin A$: Քանի որ $f_i'(x^*)$, $i \in [1 : m]$ գրադիենտները գծորեն անկախ են, ապա A-ն m չափանի ենթապարածություն է, որը փակ ուռուցիկ բազմություն է: Անջատենք այն b կետից: Համաձայն խիստ անջատման **Թեորեմ 4.1.2-ի՝** գոյություն ունեն այնպիսի p վեկտոր և $\varepsilon > 0$ թիվ, որ

$$(p, a) \leq (p, b) - \varepsilon \ \ \forall \, a \in A: \tag{5.1.4}$$

Ցույց տանք, որ

$$(p, a) = 0 \ \forall a \in A:$$

Եթե որևէ $a \in A$-ի համար $(p, a) > 0$, ապա տեղադրելով $\alpha a \in A$ $(\alpha \geq 0)$ վեկտորները (5.1.4) անհավասարության մեջ և ձգտեցնելով α-ն անվերջության՝ կստանանք հակասություն, քանի որ անհավասարության ձախ մասը կձգտի $+\infty$, իսկ աջ մասը վերջավոր թիվ է: Եթե $(p, a) < 0$, ապա կատարելով նույն դատողությունները, նորից կգանք հակասության: Այսպիսով, $(p, a) = 0 \ \forall a \in A$:
Այստեղից՝

$$(f_i'(x^*), p) = 0, \ i \in [1 : m] \Rightarrow p \in H \Rightarrow (p, b) = 0: \tag{5.1.5}$$

Բայց (5.1.4) անհավասարությունից հետևում է, որ $(p, b) >$ $> \varepsilon > 0$, ինչը հակասում է (5.1.5)-ին: Այսպիսով, ապացուցվեց, որ

$$K_M^*(x^*) = H^{\perp} = A: \blacksquare$$

Այժմ դիտարկենք հավասարության տիպի սահմանափակումներով պայմանական օպտիմիզացիայի հետևյալ խնդիրը.

$$f_0(x) \to extr,\ f_i(x) = 0,\ i \in [1:m]: \qquad (5.1.6)$$

Ենթադրվում է, որ $f_i,\ i \in [0:m]$, ֆունկցիաները անընդհատ դիֆերենցելի են R^n-ի վրա: Պահանջվում է գտնել $f_0(x)$ ֆունկցիայի էքստրեմումի կետերը

$$M \equiv \{x \in R^n /\ f_i(x) = 0,\ i \in [1:m]\}$$

բազմության վրա: Այդ կետերը կոչվում են (5.1.6) խնդրի լուծումներ: M բազմության կետերը կոչվում են (5.1.6) խնդրի թույլատրելի կետեր:
Դիցուք $\lambda = (\lambda_0, ..., \lambda_m) \in R^{m+1}$: Կազմենք Լագրանժի ֆունկցիան՝

$$L(x, \lambda) = \sum_{i=0}^{m} \lambda_i f_i(x):$$

Թեորեմ 5.1.4 *(Լագրանժի անորոշ գործակիցների մեթոդը)*: *Դիցուք x^* կետը (5.1.6) խնդրի լուծումն է:*
Այդ դեպքում գոյություն ունեն այնպիսի $\lambda_0,\ \lambda_1, \lambda_2, ..., \lambda_m$ թվեր (անորոշ գործակիցներ), որոնցից գոնե մեկը զրոյից տարբեր է, որ

$$\sum_{i=0}^{m} \lambda_i f_i'(x^*) = 0,$$

կամ որ նույն է

$$L_x'(x^*, \lambda) = 0 \Leftrightarrow L_{x_i}'(x^*, \lambda) = 0,\ i \in [1:n]: \qquad (5.1.7)$$

► Ենթադրենք, որ x^* կետը (5.1.6) խնդրում լոկալ մինիմումի կետ է (լոկալ մաքսիմումի դեպքը քննարկվում է անալոգ ձևով): Եթե $f_i'(x^*), f_2'(x^*), .., f_m'(x^*)$ գրադիենտները գծորեն անկախ են, ապա ըստ մինիմումի ընդհանուր անհրաժեշտ պայմանի և **Թեորեմ 5.1.3-ի՝** կունենանք

$$f_0'(x^*) \in K_M^*(x^*) =$$

$$= \{y / y = \sum_{i=1}^{m} \lambda_i f_i'(x^*),\ \lambda_i \in R,\ i \in [1:m]\}։$$

Այսպեղից անմիջականորեն հետևում է, որ գոյություն ունեն այնպիսի $\lambda_1, \lambda_2, ..., \lambda_m$ թվեր, որ

$$f_0'(x^*) = \sum_{i=1}^{m} \lambda_i f_i'(x^*)։$$

Այս դեպքում թեորեմի պնդումը ապացուցված է:
Եթե $f_i'(x^*), f_2'(x^*), .., f_m'(x^*)$ գրադիենտները գծորեն կախված են, ապա գոյություն կունենան գործակիցներ $\lambda_i,\ i \in [1:m]$, որոնցից գոնե մեկը զրո չէ այնպիսին, որ

$$\sum_{i=1}^{m} \lambda_i f_i'(x^*) = 0։$$

Այսպեղից հետևում է, որ

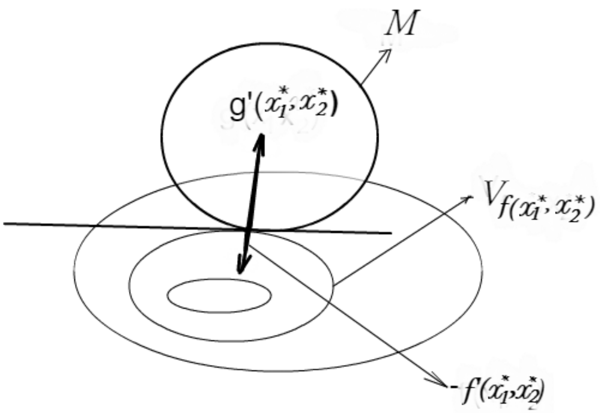
$$0 \cdot f_0'(x^*) + \sum_{i=1}^{m} \lambda_i f_i'(x^*) = 0։ \blacksquare$$

Այժմ տանք Լագրանժի անորոշ գործակիցների մեթոդի երկրաչափական մեկնաբանությունը երկչափ դեպքում: Դիցուք պետք է գտնել$f(x_1, x_2)$ ֆունկցիայի էքստրեմումի կետերը $g(x_1, x_2) = 0$ սահմանափակման դեպքում:

Ենթադրենք, որ f և g ֆունկցիաները անընդհատ դիֆերենցելի են: Հարթության վրա կառուցում ենք $M \equiv \{x \in R^2/g(x_1, x_2) = 0\}$ ողորկ կորը և f ֆունկցիայի մակարդակի V_C գծերը C պարամետրի տարբեր արժեքների դեպքում: Ակնհայտ է, որ եթե (x_1^*, x_2^*) կետը f ֆունկցիայի էքստրեմումի կետն է M բազմության վրա, ապա մակարդակի $V_{f(x_1^*, x_2^*)}$ գիծը և ողորկ M կորը (x_1^*, x_2^*) կետում շոշափում են իրար: Իրոք, եթե այդ կորերը հատվեն որևէ ոչ զրոյական անկյան տակ, ապա (x_1^*, x_2^*) կետից շարժվելով տարբեր ուղղություններով՝ մենք կհասնենք մակարդակի V_C գծերի, որոնց C պարամետրը մեծ է $f(x_1^*, x_2^*)$-ից այնպես էլ մակարդակի գծերի, որոնց C-ն փոքր է $f(x_1^*, x_2^*)$-ից: Ուրեմն (x_1^*, x_2^*)-ն էքստրեմումի կետ լինել չի կարող: Այսպիսով, $V_{f(x_1^*, x_2^*)}$ և M կորերը (x_1^*, x_2^*) կետում ունեն ընդհանուր շոշափող: Իսկ սա նշանակում է, որ $f'(x_1^*, x_2^*)$ և $g'(x_1^*, x_2^*)$ գրադիենտները պետք է լինեն զուգահեռ, քանի որ ըստ Լյուստերնիկի թեորեմի՝ մակարդակի $V_{f(x_1^*, x_2^*)}$ բազմությանը շոշափող գծի նորմալը $f'(x_1^*, x_2^*)$ գրադիենտն է, իսկ M կորինը՝ $g'(x_1^*, x_2^*)$-ն: Այսպիսով, գոյություն ունի այնպիսի λ թիվ, որ

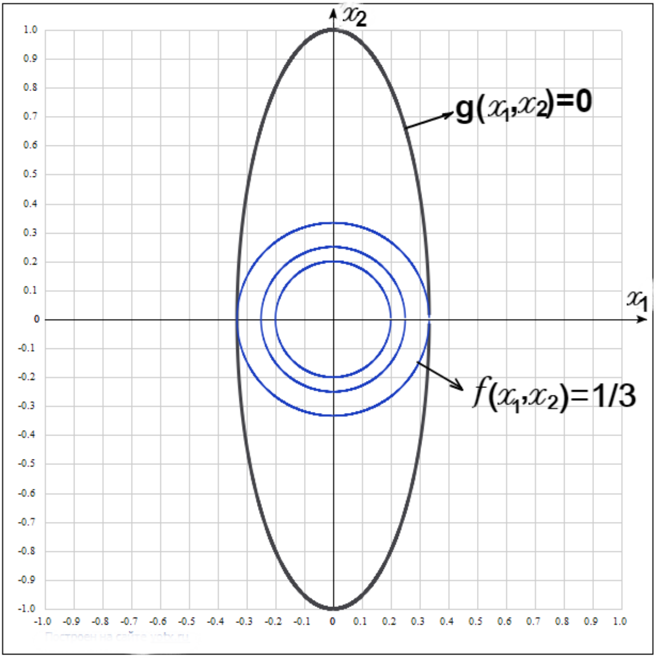
$$f'(x_1^*, x_2^*) = \lambda g'(x_1^*, x_2^*),$$

ինչը նշանակում է $f'(x_1^*, x_2^*)$ և $g'(x_1^*, x_2^*)$ գրադիենտների գծային կախվածություն (տես գծ.5.5):



Գծ. 5.5: Լագրանժի անորոշ գործակիցների մեթոդի երկրաչափական մեկնաբանությունը

Օրինակ: Գտնենք $f(x_1, x_2) = x_1^2 + x_2^2$ ֆունկցիայի մինիմումի կետերը $g(x_1, x_2) = 9x_1^2 + x_2^2 = 0$ սահմանափակման դեպքում: Արդեն գիտենք, որ $g(x_1, x_2) = 0$ բազմությունը հարթության վրա էլիպսներ են, իսկ ֆունկցիայի մակարդակի $V_C = \{x \in R^n / f(x_1, x_2) = C\}$ բազմությունները $(0, 0)$ կենտրոնով $\sqrt{C}$ շրջանագծեր են: Հետևաբար C-ի այն արժեքը, որի դեպքում շրջանագիծը կշոշափի էլիպսին, դա կլինի f ֆունկցիայի մինիմալ արժեքը, իսկ շոշափման կետերը կլինեն նրա մինիմումի կետեր: Գծ.5.6-ից երևում է, որ $C = 1/3$, իսկ մինիմումի կետերն են՝ $(-1/3, 0)$, $(-1/3, 0)$:



Գծ. 5.6: $f(x_1, x_2) = x_1^2 + x_2^2$ ֆունկցիայի մինիմումը $g(x_1, x_2) = 9x_1^2 + x_2^2 = 0$ բազմության վրա: Գծապատկերային լուծումը

Օրինակ: Գտնել $A(\xi_1, \xi_2)$ կետի պրոյեկցիան

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$$

էլիպսի վրա: Դա նշանակում է, որ պետք է լուծել հետևյալ խնդիրը.

$$f_0(x_1, x_2) = (x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 \longrightarrow min,$$

$$f_1(x_1, x_2) = \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - 1 = 0:$$

Կազմենք Լագրանժի ֆունկցիան

$$L = \lambda_0 f_1(x_1, x_2) + \lambda_2 f_2(x_1, x_2):$$

Գրենք մինիմումի անհրաժեշտ պայմանը.

$$L_{x_i} = 0 \Leftrightarrow \lambda_0(x_i - \xi_i) + \lambda_1 \frac{x_i}{a_i^2},\ i = 1, 2:$$

Եթե $\lambda_0 = 0$, ապա $\lambda_1 \neq 0$, քանի որ Լագրանժի գործակիցներից գոնե մեկը պետք է լինի զրոյից տարբեր: Այդ դեպքում ստանում ենք, որ $x_1 = x_2 = 0$: Բայց $(0, 0)$ կետը էլիպսի վրա չէ: Հետևաբար $\lambda_0 = 0$ դեպքը հնարավոր չէ: Ուստի կարող ենք ընդունել, որ $\lambda_0 = 1$: Այստեղից՝

$$x_i = \xi_i \frac{a_i^2}{a_i^2 + \lambda},\ i = 1, 2:$$

Այժմ տեղադրելով այդ արժեքները էլիպսի հավասարման մեջ՝ λ-ի նկատմամբ ստանում ենք

$$\varphi(\lambda) = \frac{\xi_1^2 a_1^2}{(a_1^2 + \lambda)^2} + \frac{\xi_2^2 a_2^2}{(a_2^2 + \lambda)^2} = 1$$

հավասարումը.

Խնդրի ստացիոնար կետերի քանակը հավասար է այս հավասարման արմատների քանակին, որը չի գերազանցում չորսի: Դիցուք կետը գտնվում է էլիպսից դուրս: Այդ դեպքում

$$\varphi(0) = \frac{\xi_1^2 a_1^2}{a_1^2} + \frac{\xi_2^2 a_2^2}{a_2^2} > 1$$

և խնդիրն ունի երկու ստացիոնար կետ (տե՛ս գծ.5.7):

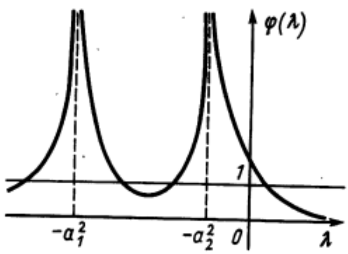
Քանի որ էլիպսը փակ սահմանափակ բազմություն է, ապա համաձայն Վայերշտրասի թեորեմի՝ խնդիրը ունի լուծում: Անհրաժեշտ է լուծել $\varphi(\lambda) = 1$ հավասարումը, գտնել ստացիոնար կետերը և տեղադրելով դրանք f_0 ֆունկցիայի մեջ՝ գտնել այդ թվերից ամենափոքրը:

$$x_i - \xi_i + \frac{\lambda_1 x_i}{a_i^2} = 0,\ i = 1,2$$

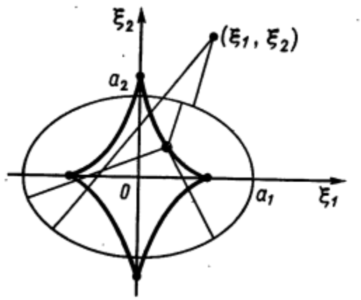
հավասարությունները ունեն պարզ երկրաչափական մեկնաբանություն. $\xi - x$ վեկտորը էլիպսի նորմալ վեկտորն է x կետում: Այժմ նկարագրենք հարթության այն կետերը, որոնք ունեն մեկ, երկու, երեք կամ չորս նորմալ էլիպսի վրա: Դրա համար գտնենք $\varphi'(\lambda) = 0$ հավասարման արմատները և նրանք տեղադրելով $\varphi(\lambda) = 1$ հավասարման մեջ՝ կստանանք

$$(\xi_1 a_1)^{2/3} + (\xi_2 a_2)^{2/3} = (a_1^2 - a_2^2)^{2/3}:$$

Այս հավասարումը աստրոիդի հավասարում է: Աստրոիդից դուրս յուրաքանչյուր կետ ունի երկու նորմալ, ներսի կետերը՝ չորս, իսկ աստրոիդի վրայի կետերը, բացի գագաթներից, ունեն երեք նորմալ էլիպսի վրա (տես գծ.5.8):



Գծ. 5.7:



Գծ. 5.8:

Այժմ ենթադրենք, որ (5.1.6) խնդրում բոլոր ֆունկցիաները երկու անգամ անընդհատ դիֆերենցելի են: Ճիշտ է հետևյալ պնդումը (տե՛ս, օրինակ՝ [5]):

Թեորեմ 5.1.5 *(Էքստրեմումի երկրորդ կարգի անհրաժեշտ պայմանը)*: *Դիցուք x^*-ը (5.1.6) խնդրում լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է և $f_1'(x^*),\ f_2'(x^*), ..., f_m'(x^*)$ վեկտորները գծորեն անկախ են: Այդ դեպքում գոյություն ունեն այնպիսի $\lambda_0 = 1,\ \lambda_1, ..., \lambda_m$ թվեր, որ*

$$(L''_{xx}(x^*, \lambda)h, h) \geq 0 \ \ ((L''_{xx}(x^*, \lambda)h, h) \leq 0)\ \forall h \in H =$$

$$= \{h \in R^n / (f_i'(x^*), h) = 0,\ i \in [1:m]\} :$$

Թեորեմ 5.1.6 *(Երկրորդ կարգի բավարար պայմանը)*: *Դիցուք x^*-ը (5.1.6) խնդրի թույլատրելի կետ է և գոյություն ունի այնպիսի $\lambda \in R^{m+1}$ վեկտոր, որ տեղի ունեն հետևյալ պայմանները.*

1) $\lambda_0 = 1,$

2) $L'_{x_i}(x^*, \lambda) = 0,\ i \in [1:n]$,

3) *կամայական ոչ զրոյական* $h \in H$ *վեկտորի համար տեղի ունի*

$$(L''_{xx}(x^*, \lambda)h, h) > 0 \ ((L''_{xx}(x^*, \lambda)h, h) < 0) \tag{5.1.8}$$

անհավասարությունը:

Այդ դեպքում x^**-ը* (5.1.6) *խնդրում լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է:*

► Եթե x^* կետը M բազմության առանձնացված կետ է, ապա թեորեմի եզրակացությունը փրիվիալ է: Դիցուք այժմ x^*-ը M բազմության սահմանային կետ է և այն խնդրում լոկալ մինիմումի կետ չէ: Այդ դեպքում գոյություն կունենա այնպիսի $\{x^k\}$ հաջորդականություն, որը բավարարում է հետևյալ պայմաններին.

$$x^k \in M,\ x^k \to x^*,\ f_0(x^k) < f_0(x^*): \tag{5.1.9}$$

x^k -ն ներկայացնենք հետևյալ տեսքով.

$$x^k = x^* + \alpha_k h^k, \text{ որտեղ } h^k = (x^k - x^*)/\alpha_k:$$

Քանի որ, $\|h^k\| = 1$, ապա ընդհանրությունը չսահմանափակելով կարող ենք ենթադրել, որ

$$h^k \to h \neq 0:$$

Հաշվի առնելով (5.1.9) պայմանը՝ ունենք

$$0 = f_i(x^k) - f_i(x^*) = (f_i'(x^*), \alpha_k h^k) + o(\alpha_k),\ i \in [1:m]:$$

Բաժանելով այս առնչությունները α_k-ի վրա և անցնելով սահմանի՝ կստանանք

$$(f_i'(x^*), h) = 0,\ i \in [1:m]:$$

Այսինքն՝ $h \in H$: Քանի որ ըստ ենթադրության $\lambda_0 = 1$, ապա (5.1.9)-ից ստանում ենք

$$L(x^k, \lambda) = f_0(x^k) + \sum_{i=1}^{m} \lambda_i f_i(x^k) \leq f_0(x^k) \leq$$

$$\leq f_0(x^*) = f_0(x^*) + \sum_{i=1}^{m} \lambda_i f_i(x^*) = L(x^*, \lambda): \qquad (5.1.10)$$

Թեորեմի պայմաններից հետևում է, որ $L(x, \lambda)$ Լագրանժի ֆունկցիան երկու անգամ դիֆերենցելի է x^* կետում: Հետևաբար ներկայացնելով այդ ֆունկցիան ըստ Թեյլորի բանաձևի՝ x^* կետում ստանում ենք

$$L(x^k, \lambda) = L(x^*, \lambda) + (L'_x(x^*, \lambda), \alpha_k h^k) +$$

$$+ \frac{1}{2}(L''_{xx}(x^*, \lambda)(\alpha_k h^k), \alpha_k h^k) + o(\alpha_k^2):$$

Քանի որ, ըստ ենթադրության, $L'_x(x^*, \lambda) = 0$, ապա այստեղից և (5.1.10) անհավասարությունից հետևում է, որ

$$\frac{\alpha_k^2}{2}(L''_{xx}(x^*, \lambda)h^k, h^k)) + o(\alpha_k^2) \leq 0 :$$

Այս անհավասարության երկու մասերը բաժանելով α_k^2 թվի վրա և անցնելով սահմանի՝ կստանանք

$$(L''_{xx}(x^*, \lambda)h, h) \leq 0,$$

որը հակասում է թեորեմի (5.1.8) պայմանին: ■

Պարզագույն դեպքերում Լագրանժի անորոշ գործակիցների մեթոդը թույլ է տալիս բացահայտ տեսքով գտնել մաթեմատիկական ծրագրավորման խնդիրների լուծումները հավասարությունների տիպի սահմանափակումների դեպքում: Դրա համար պետք է կատարել է հետևյալ քայլերը.

- Կազմել Լագրանժի ֆունկցիան:
- Գրել էքստրեմումի առաջին կարգի անհրաժեշտ պայմանը Լագրանժի ֆունկցիայի համար և ստանալ հավասարումների համակարգ, որով բնութագրվում է խնդրի ստացիոնար կետերի բազմությունը:
- Գտնել ստացված համակարգի լուծումները:
- Էքստրեմումի երկրորդ կարգի բավարար պայմանների միջոցով այդ լուծումներից անջատել էքստրեմումի կետերը:
  Այս ալգորիթմը մեկնաբանենք օրինակներով:

Օրինակ 1: Լուծել հետևյալ խնդիրը.

$$f_0(x) = x_1^2 + x_2^2 \to extr,$$

$$f_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0:$$

Լուծում: Կազմենք Լագրանժի ֆունկցիան

$$L(x, \lambda) = \lambda_0 f_0(x) + \lambda_1 f_1(x) =$$

$$= \lambda_0(x_1^2 + x_2^2) + \lambda_1((x_1 - 1)^2 + x_2^2 - 4):$$

Ըստ (5.1.7) անհրաժեշտ պայմանի՝ ունենք

$$\left\{ \begin{array}{l} L'_{x_1}(x, \lambda_0, \lambda_1) = 2\lambda_0 x_1 + 2\lambda_1(x_1 - 1) = 0, \\ L'_{x_2}(x, \lambda_0, \lambda_1) = 2\lambda_0 x_2 + 2\lambda_1 x_2 = 0, \\ f_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0 : \end{array} \right. \tag{5.1.11}$$

Եթե, $\lambda_0 = 0$, ապա (5.1.11) համակարգից ստանում ենք

$$2\lambda_1(x_1 - 1) = 0,\ 2\lambda_1 x_2 = 0,\ (x_1 - 1)^2 + x_2^2 - 4 = 0: \tag{5.1.12}$$

Քանի որ λ_0 և λ_1 գործակիցները միաժամանակ զրո չեն, ապա $\lambda_1 \neq 0$:

Հետևաբար (5.1.12) համակարգի առաջին երկու պայմաններից կստանանք

$$x_1 = 1,\ x_2 = 0,$$

որը չի բավարարում երրորդ հավասարմանը: Այսպիսով, $\lambda_0 \neq 0$: Ընդհանրությունը չսահմանափակելով, կարող ենք ենթադրել, որ $\lambda_0 = 1$: Այդ դեպքում (5.1.11) համակարգը կընդունի հետևյալ տեսքը.

$$\begin{cases} 2x_1 + 2\lambda_1(x_1 - 1) = 0, \\ 2x_2 + 2\lambda_1 x_2 = 0, \\ (x_1 - 1)^2 + x_2 - 4 = 0 : \end{cases} \tag{5.1.13}$$

Դիտարկենք այս համակարգի երկրորդ հավասարումը: Եթե $x_2 = 0$, ապա երրորդից կստանանք $x_1 = 3,\ x_1 = -1$, իսկ առաջին հավասարումից՝ $\lambda_1 = -3/2$:

Եթե $x_2 \neq 0$, ապա երկրորդից կունենանք $\lambda_1 = -1$:

Այդ դեպքում առաջին հավասարումը տեղի չունի, այսինքն՝ (5.1.13) համակարգը համատեղելի չէ:
Այսպիսով, ստանում ենք երկու ստացիոնար կետեր՝

$$x_1^* = 3,\ x_2^* = 0,\ \lambda_1 = -\tfrac{3}{2};$$
$$x_1^* = -1,\ x_2^* = 0,\ \lambda_1 = -\tfrac{1}{2}:$$

Ստուգենք երկրորդ կարգի (5.1.8) բավարար պայմանները այդ կետերի համար: Ունենք՝

$$(L''_{xx}h, h) = 2(1 + \lambda_1)h_1^2 + 2(1 + \lambda_1)h_2^2,$$

$$h \in H = \{h = (h_1, h_2) \in R^2 / 2(x_1^* - 1)h_1 + 2x_2^* h_2 = 0:$$

Այսպեղից հեշտ է պեսնել, որ $A(3,0)$ կետի համար տեղի ունի

$$(L''_{xx}h,h) < 0 \; \forall h \in H, \; h \neq 0$$

անհավասարությունը: Ուստի A-ն լոկալ մաքսիմումի կետ է: Նման ձևով համոզվում ենք, որ $B(-1,0)$-ն լոկալ մինիմումի կետ է: Մյուս կողմից, քանի որ f_0 ֆունկցիան հասնում է իր մեծագույն ու փոքրագույն արժեքներին, ապա B կետը գլոբալ մինիմումի կետ է, իսկ A-ն գլոբալ մաքսիմումի կետ է:

Օրինակ 2: Լուծել հետևյալ խնդիրը.

$$\begin{aligned} &f_0(x) = x_1^2 + x_2^2 + x_3^2 \to min, \\ &f_1(x) = x_1^2 + x_2^2 - x_3 = 0, \\ &f_2(x) = x_1 + x_2 + x_3 - 4 = 0: \end{aligned}$$

Ըստ (5.1.7) անհրաժեշտ պայմանի՝ ունենք

$$\begin{cases} L'_{x_1}(x,\lambda) = 2\lambda_0 x_1 + 2\lambda_1 x_1 + \lambda_2 = 0, \\ L'_{x_2}(x,\lambda) = 2\lambda_0 x_2 + 2\lambda_1 x_2 + \lambda_2 = 0, \\ L'_{x_3}(x,\lambda) = 2\lambda_0 x_3 - \lambda_1 + \lambda_2 = 0, \\ f_1(x) = x_1^2 + x_2^2 - x_3 = 0, \\ f_2(x) = x_1 + x_2 + x_3 - 4 = 0: \end{cases}$$

Եթե $\lambda_0 = 0$, ապա կստանանք հավասարումների հետևյալ համակարգը.

$$\begin{cases} 2\lambda_1 x_1 + \lambda_2 = 0, \\ 2\lambda_1 x_2 + \lambda_2 = 0, \\ -\lambda_1 + \lambda_2 = 0, \\ x_1^2 + x_2^2 - x_3 = 0, \\ x_1 + x_2 + x_3 - 3 = 0: \end{cases} \tag{5.1.14}$$

Եթե $\lambda_1 = 0$, ապա (5.1.14) համակարգի երրորդ հավասարումից հետևում է, որ $\lambda_2 = 0$, որը հնարավոր չէ, որովհետև բոլոր գործակիցները միաժամանակ զրո չեն: Եթե $\lambda_1 \neq 0$, ապա (5.1.14) համակարգի առաջին երեք հավասարումներից սպանում ենք

$$x_1 = x_2 = -\frac{1}{2}:$$

Տեղադրելով այս արժեքները համակարգի վերջին երկու հավասարումների մեջ՝ սպանում ենք հակասություն: Այսպիսով, $\lambda_0 \neq 0$: Ընդունենք $\lambda_0 = 1$: Այդ դեպքում համակարգը կընդունի հետևյալ տեսքը.

$$\begin{cases} 2x_1(1+\lambda_1) + \lambda_2 = 0, \\ 2x_2(1+\lambda_1) + \lambda_2 = 0, \\ 2x_3 - \lambda_1 + \lambda_2 = 0, \\ x_1^2 + x_2^2 - x_3 = 0, \\ x_1 + x_2 + x_3 - 4 = 0: \end{cases} \tag{5.1.15}$$

(5.1.15) համակարգը ունի հետևյալ երկու լուծումները.

$$x_1^* = 1,\ x_2^* = 1,\ x_3^* = 2, \lambda_1 = \tfrac{2}{3},\ \lambda_2 = -\tfrac{10}{3};$$
$$x_1^* = -2,\ x_2^* = -2,\ x_3^* = 8,\ \lambda_1 = -\tfrac{20}{3}, \lambda_2 = -\tfrac{68}{3}:$$

Ստուգենք երկրորդ կարգի (5.1.8) բավարար պայմանները $A(1,1,2)$ և $B(-2,-2,8)$ կետերի համար:
Ունենք

$$(L''_{xx}h, h) = 2(1+\lambda_1)h_1^2 + 2(1+\lambda_1)h_2^2 + 2h_2h_3,$$

$$h \in H = \{h \in R^3 / 2x_1^* h_1 + 2x_2^* h_2 - h_3 = 0,$$
$$h_1 + h_2 + h_3 = 0\}:$$

Այստեղից $A(1,1,2)$ կետի համար կստանանք

$$h_1 = -h_2, h_3 = 0 \Rightarrow (L''_{xx}h, h) = \frac{10}{3}h_2^2 > 0\ \forall h \neq 0:$$

Այսինքն՝ A-ն լոկալ մինիմումի կետ է: Նույն ձևով ստանում ենք, որ $B(-2,-2,8)$-ն լոկալ մաքսիմումի կետ է:

ԽՆԴԻՐՆԵՐ

1. Գրաֆիկորեն լուծել հետևյալ խնդիրը.

$$\begin{aligned} & x_1 + x_2^2 \to min, \\ & x_1^2 + x_2^2 - 8 = 0: \end{aligned}$$

2. Գրաֆիկորեն լուծել հետևյալ խնդիրը.

$$\begin{aligned} & x_1^2 + x_2^2 \to min, \\ & x_1^2 + 2x_2^2 - 8 = 0: \end{aligned}$$

3. Ստուգել, արդյո՞ք $(0,2)$ կետը հետևյալ խնդրի լուծումն է, թե ոչ:

$$\begin{aligned} & x_1^2 + x_2^2 \to min, \\ & x_2 + x_1^2 - 2 = 0: \end{aligned}$$

4. Ստուգել, արդյո՞ք $(-2,2)$ կետը հետևյալ խնդրի լուծումն է, թե ոչ:

$$\begin{aligned} & x_1 x_2 \to min, \\ & x_1^2 + x_2^2 - 8 = 0: \end{aligned}$$

5. Լագրանժի գործակիցներով լուծել հետևյալ խնդիրը.

$$\begin{aligned} & 2x_1^2 + x_2^2 + 2x_3^2 \to extr, \\ & 2x_1 + x_2 + 2x_3 = 12, \\ & x_1 + 2x_2 + x_3 = 10): \end{aligned}$$

6. Լագրանժի գործակիցներով լուծել հետևյալ խնդիրը.

$$x_1^2 - x_2^2 + x_3^2 \to max,$$
$$x_1 + 2x_2^2 - x_3 = 4,$$
$$2x_1 + 2x_2 + x_3 = 14:$$

5.2 Լագրանժի անորոշ գործակիցների մեթոդը (խառը սահմանափակումների դեպքը)

Այժմ դիտարկենք պայմանական օպտիմիզացիայի ընդհանուր խնդիրը, որտեղ սահմանափակումները տրվում են հավասարություններով և անհավասարություններով՝

$$f_0(x) \to extr,$$

$$f_i(x) = 0,\ i \in [1:k],\ f_i(x) \leq 0,\ i \in [k+1:m]: \quad (5.2.1)$$

Այստեղ ենթադրվում է, որ $f_i(x),\ i \in [0:m]$ ֆունկցիաները անընդհատ դիֆերենցելի են R^n-ի վրա:
$x \in R^n$ կետը կոչվում է թույլատրելի, եթե այն բավարարում է (5.2.1) խնդրի բոլոր սահմանափակումներին: $x^* \in R^n$ կետը կոչվում է (5.2.1) խնդրի լուծում, եթե այն $f_0(x)$ ֆունկցիայի էքստրեմումի կետ է

$$M \equiv \{x \in R^n /\ f_i(x) = 0,\ i \in [1:k],$$

$$f_i(x) \leq 0,\ i \in [k+1:m]\}$$

բազմության վրա:

Հիշենք նաև ակտիվ-պասիվ սահմանափակումների սահմանումը, որը ինչպես տեսանք էական դեր խաղաց ուռուցիկ խնդիրներում մինիմումի կետի որոշման գրադիենտի պրոյեկտման ընթացակարգում:

Դիցուք $\bar{x}$-ը թույլատրելի կետ է (5.2.1) խնդրում: $f_i(x) \leq 0$ $(i \in [k+1:m])$ սահմանափակումը կոչվում է ակտիվ այդ կետում, եթե $f_i(\bar{x}) = 0$: Եթե $f_i(\bar{x}) < 0$, ապա այդ սահմանա–փակումը կոչվում է պասիվ: $I_{\text{ա}}(\bar{x})$ սիմվոլով նշանակենք $\bar{x}$ կետում ակտիվ սահմանափակումների ինդեքսների բազմությունը.

$$I_{\text{ա}}(\bar{x}) = \{i \in [k+1, m] / \ f_i(\bar{x}) = 0\}:$$

Ճիշտ են հետևյալ պնդումները (տե՛ս, օրինակ՝ [20]):

Թեորեմ 5.2.1 *(Էքստրեմումի առաջին կարգի անհրա–ժեշտ պայմանը)*: *Դիցուք x^* կետը (5.2.1) խնդրում լոկալ մինիմումի (լոկալ մաքսիմումի) կետ է: Այդ դեպքում գոյություն ունեն այնպիսի $\lambda_0, ..., \lambda_m$ թվեր, որոնցից գոնե մեկը զրոյից տարբեր է, որ*

ա) $L'_{x_i}(x^*) = 0, \ i \in [1:n]$, *որտեղ*

$$L(x, \lambda) = \sum_{i=0}^{m} \lambda_i f_i(x),$$

բ) $\lambda_i \geq 0 \ (\lambda_i \leq 0), \ i \in [k+1:m],$

գ) $\lambda_i f_i(x^*) = 0, \ i \in [k+1:m]$*:*

գ) հավասարությունը կոչվում է պասիվ-ակտիվ սահմա–նափակումների պայման:

Հետևանք 5.2.1: *Եթե* $f_0, f_i, \ \ i = k+1, \ldots, m$ *ֆունկ–ցիաները ուռուցիկ են, իսկ* $f_i, i = 1, \ldots, k$ *ֆունկցիաները՝*

գծային, ապա թեորեմ 5.2.1 պայմանները հանդիսանում են գլոբալ մինիմումի բավարար պայմաններ:

Թեորեմ 5.2.2 *(Առաջին կարգի բավարար պայմանը)*: *Դիցուք x^* վեկտորը բավարարում է հետև– յալ պայմաններին.*

1) x^*- *ը* (5.2.1) *խնդրի թույլատրելի կետ է,*
2) *գոյություն ունի այնպիսի* $\lambda = (\lambda_0, \lambda_1, ..., \lambda_m)$ $\lambda_0 = 1$ *վեկտոր, որ* (x^*, λ) *զույգը բավարարում է թեորեմ* 5.2.1-*ի ա), բ), գ) պայմաններին,*
3) x^* *կետում* (5.2.1) *խնդրի ակտիվ սահմանա– փակումների և հավասարությունների քանակ– ների գումարը հավասար է* n-*ի:*

Այդ դեպքում, եթե $\lambda_i > 0$, $i \in I_{ա}(x^*)$, *ապա* x^*-*ը* (5.2.1) *խնդրում լոկալ մինիմումի կետ է, եթե* $\lambda_i < 0$, $i \in I_{ա}(x^*)$, *ապա* x^*-*ը* (5.2.1) *խնդրում լոկալ մաքսիմումի կետ է:*

Այժմ ձևակերպենք քայլերի այն հերթականությունը, որոնց միջոցով կարելի է լուծել խառը սահմանափակումներով մաթեմատիկական ծրագրավորման խնդիրները:

- Կազմել Լագրանժի ֆունկցիան:
- Գրել էքստրեմումի առաջին կարգի անհրաժեշտ պայմանը Լագրանժի ֆունկցիայի համար և ստանալ հավասարումների համակարգ, որով բնութագրվում է խնդրի ստացիոնար կետերի բազմությունը:
- Գրել պասիվ-ակտիվ սահմանափակումների պայմանները և անհավասարություններին հա– մապատասխանող Լագրանժի գործակիցների ոչ բացասական (ոչ դրական) լինելու պայմանները:

- Լուծել ստացված համակարգերը՝ հաշվի առնելով Լագրանժի գործակիցների նշանները:
- Օպտիմալության առաջին կարգի բավարար պայմանների միջոցով այդ լուծումներից անջատել էքստրեմումի կետերը:
  Այս ալգորիթմը մեկնաբանենք օրինակի միջոցով:

Օրինակ: Լուծել հետևյալ խնդիրը.

$$\begin{aligned} & x_1 - x_2^2 \to extr, \\ & x_1 - x_2 - 1 = 0, \\ & x_1^2 + x_2^2 - 5 \leq 0 : \end{aligned}$$

Կազմենք Լագրանժի ֆունկցիան

$$L(x, \lambda) = \lambda_0(x_1 - x_2^2) + \lambda_1(x_1 - x_2 - 1) + \lambda_2(x_1^2 + x_2^2 - 5):$$

Ըստ Էքստրեմումի անհրաժեշտ պայմանի՝ ունենք հավասարումների և անհավասարումների հետևյալ համակարգը.

(ա) $L'_{x_1}(x, \lambda) = \lambda_0 + \lambda_1 + 2\lambda_2 x_1 = 0,$

(բ) $L'_{x_2}(x, \lambda) = -2\lambda_0 x_2 - \lambda_1 + 2\lambda_2 x_2 = 0,$

(գ) $x_1 - x_2 - 1 = 0,\ x_1^2 + x_2^2 - 5 \leq 0,$

(դ) $\lambda_2 \geq 0,$

(ե) $\lambda_2(x_1^2 + x_2^2 - 5) = 0:$

Դիտարկենք երկու դեպք.

1) $\lambda_0 = 0,$

2) $\lambda_0 \neq 0:$

Առաջին դեպքում համակարգի (ա) և (բ) հավասարումներից կստանանք

$$\begin{cases} \lambda_1 + 2\lambda_2 x_1 = 0, \\ -\lambda_1 + 2\lambda_2 x_2 = 0 \end{cases} \tag{5.2.2}$$

համակարգը: Եթե $\lambda_2 = 0$, ապա (5.2.2) համակարգից կստանանք $\lambda_1 = 0$, այսինքն՝ բոլոր գործակիցները զրո են, որը հակասություն է:
Եթե $\lambda_2 \neq 0$, ապա (գ) և (ե) պայմաններից կստանանք

$$\begin{cases} x_1^2 + x_2^2 - 5 = 0, \\ x_1 - x_2 - 1 = 0 \end{cases}$$

համակարգը: Այս համակարգը ունի երկու լուծում.

$$\begin{cases} x_1 = 2, \\ x_2 = 1 \end{cases} \tag{5.2.3}$$

կամ

$$\begin{cases} x_1 = -1, \\ x_2 = -2: \end{cases} \tag{5.2.4}$$

Գումարելով (5.2.2) համակարգի հավասարումները՝ կստանանք

$$2\lambda_2(x_1 + x_2) \Rightarrow x_1 = -x_2,$$

որը հակասում է (5.2.3) և (5.2.4) համակարգերին: Հետևաբար առաջին դեպքը հնարավոր չէ:
Դիտարկենք երկրորդ դեպքը. $\lambda_0 \neq 0$: Ընդունելով $\lambda_0 = 1$՝ (ա) և (բ) պայմաններից կստանանք

$$\begin{cases} 1 + \lambda_1 + 2\lambda_2 x_1 = 0, \\ -2x_2 - \lambda_1 + 2\lambda_2 x_2 = 0: \end{cases} \tag{5.2.5}$$

Եթե $\lambda_2 = 0$, ապա (5.2.5) համակարգից և (գ) հավասարությունից կստանանք

$$\begin{cases} 1+\lambda_1 = 0, \\ -2x_2 - \lambda_1 = 0, \\ x_1 - x_2 - 1 = 0\text{:} \end{cases}$$

Լուծելով այս համակարգը՝ ստանում ենք ստացիոնար հետևյալ կետը.

$$A(3/2, 1/2),\ \lambda_1 = -1,\ \lambda_2 = 0\text{:}$$

$\lambda_2 \neq 0$ դեպքում ունենք

$$\begin{cases} x_1^2 + x_2^2 - 5 = 0, \\ x_1 - x_2 - 1 = 0, \\ 1 + \lambda_1 + 2\lambda_2 x_1 = 0, \\ -2x_2 - \lambda_1 + 2_2 x_2 = 0 \end{cases}$$

համակարգը: Լուծելով այս համակարգը՝ ստանում ենք ևս երկու ստացիոնար կետեր.

$$B(2,1),\ \lambda_1 = -5/3,\ \lambda_2 = 1/6,$$
$$C(-1,-2),\ \lambda_1 = 2/3,\ \lambda_2 = 5/6\text{:}$$

Այս երեք կետերի համար ստուգենք էքստրեմումի առաջին կարգի բավարար պայմանները:

B կետի համար ակտիվ է $x_1^2 + x_2^2 - 5 \leq 0$ սահմանափակումը և նրան համապատասխան Լագրանժի գործակիցը դրական է: Մյուս կողմից, ակտիվ սահմանափակումների և հավասարությունների քանակների գումարը հավասար է երկուսի, որը անհայտների թիվն է: Նշանակում է B-ն լոկալ մինիմումի կետ է: Նույն ձևով՝ C-ն լոկալ մինիմումի կետ է: Բայց քանի որ խնդրի սահմանափակումների բազմու– թյունը կոմպակտ է, ապա նպատակային ֆունկցիան ունի գլոբալ մինիմումի և մաքսիմումի կետեր: Հաշվելով ստացված կետերում նպատակային ֆունկցիայի արժեքները՝ պարզում ենք, որ A կետը գլոբալ մաքսիմումի կետ է, C-ն գլոբալ մինիմումի կետ է, իսկ B-ն լոկալ մինիմումի կետ է:

ԽՆԴԻՐՆԵՐ

1. Լուծել խառը սահմանափակումներով հետևյալ խնդիրը.
   $x_1^2 - x_2 \to min,$
   $x_1 + x_2 - 6 = 0,$
   $1 - x_1 \leq 0,$
   $x_1^2 + x_2^2 - 26 \leq 0$:
2. Լուծել խառը սահմանափակումներով հետևյալ խնդիրը.
   $x_1^2 + x_2^2 + x_3^2 \to min,$
   $2x_1 - x_2 + x_3 \leq 5,$
   $x_1 + x_2 + x_3 = 3$:
3. Գտնել շոշափող $K_M(x)$ կոնը M բազմության համար x կետում:

ա) $M = \{(x_1, x_2) \in R^2 \backslash int(R_+^2) : x_1^2 + x_2^2 \leq 1\},$
$x = (0, 1)$:

բ) $M = \{(x_1, x_2) \in R^2 \backslash int(R_+^2) : x_1^2 + x_2^2 \leq 1\},$
$x = (0, 0)$:

գ) $M=\{(x_1,x_2)/x_1^2\le x_2^3\},\ x=(0,0)$:

դ) $M=\{0,1,\dfrac{1}{2},...,\dfrac{1}{n},...\},\ x=0$:

Ցուցում: Դիցուք

$$\varphi(\alpha)=\begin{cases}-\alpha+\frac{1}{n},\ \alpha\in(\frac{1}{2}(\frac{1}{n}+\frac{1}{n+1})\frac{1}{n}]\\ -\alpha+\frac{1}{n+1},\ \alpha\in(\frac{1}{2}(\frac{1}{n}+\frac{1}{n+1})\frac{1}{n+1}]:\end{cases}$$

Ապացուցել, որ $\alpha+\varphi(\alpha)\in M,\ \ \varphi(\alpha)=o(\alpha)$, այսինքն՝ $h=1$ վեկտորը շոշափող է: Ուրեմն $K_M(0)=R_+$:

ե) $M=\{0,1,\dfrac{1}{2},...,\dfrac{1}{2^n},...\},\ x=0$:

4. Ապացուցել հետևյալ պնդումը. որպեսզի $K\subseteq R^n$ կոնը լինի ուռուցիկ, անհրաժեշտ է և բավարար, որ $\forall x,\ y\in K\Rightarrow x+y\in K$:

5*. Դիցուք $K\subseteq R^n$ փակ ուռուցիկ կոն է: Ապացուցել, որ $K^{**}=K$:

6. Դիցուք $K_1,\ K_2\subseteq R^n$-ը փակ ուռուցիկ կոներ են: Ապացուցել, որ

ա) $(K_1+K_2)^*=K_1^*\cap K_2^*$;

բ) $(K_1\cap K_2)^*=\overline{K_1^*+K_2^*}$:

Լուծում: Ունենք

$$(K_1\cap K_2)^*=(K_1^{**}\cap K_2^{**})^*=$$

$$=((K_1^*+K_2^*)^*)^*=(K_1^*+K_2^*)^{**}=\overline{K_1^*+K_2^*}:$$

Գլուխ 6

Վարիացիոն հաշիվ

Վարիացիոն հաշիվը օպտիմիզացիայի բաժին է, որտեղ ուսումնասիրվում են ինտեգրալային տիպի ֆունկցիոնալների մինիմիզացիայի խնդիրները որոշ ֆունկցիոնալ տարածություններում: Լեոնարդ Էյլերի ֆունդամենտալ աշխատանքների շնորհիվ վարիացիոն հաշվը դառնում է ինքնուրույն մաթեմատիկական առարկա և տրվում է այդ դասի խնդիրների լուծման ընդհանուր մեթոդիկա: Այդպիսի խնդիրներում էքստրեմալների բնութագրման համար տրվեց ընդհանուր անհրաժեշտ պայման: Վարիացիոն հաշվի խնդիրների լուծման մեթոդների մշակման գործում իրենց կարևոր ներդրումները ունեցան նաև սովետական մաթեմատիկոսներ Մ. Ա. Լավրենտևը, Ս. Լ. Սոբոլևը, Ն. Ն. Բոգոլյուբովը և այլոք (տես [6, 25]):

Այս գլխում դիտարկվում է

$$I(y) = \int_{x_0}^{x_1} L(x, y, y')\, dx$$

տիպի ֆունկցիոնալի մինիմիզացիայի (մաքսիմիզացիայի) խնդիրը $C^1[x_0, x_1]$ ֆունկցիոնալ տարածության վրա: Ցույց է տրվում, որ այդ ֆունկցիոնալի էքստրեմումները բավարարում են եզրային պայմաններով երկրորդ կարգի մի դիֆերենցիալ հավասարմանը, որը կոչվում է Էյլերի հավասարում: Այդ հավասարումով կապ է ստեղծվում վարիացիոն հաշվի և դիֆերենցիալ հավասարումների տեսությունների միջև: Այդ իսկ պատճառով վարիացիոն հաշվի մեթոդները օգտագործում են եզրային պայմաններով դիֆերենցիալ հավասարումների լուծումների գոյության ապացույցներում: Այս գլխում վարիացիոն հաշվի պարզագույն խնդիրների էքստրեմալների բնորոշման համար տրվում են որոշ բավարար պայմաններ: Վարիացիոն հաշվի տեսության ավելի խորը ուսումնասիրությունների հետ կարելի է ծանոթանալ [3, 4, 31, 32] աշխատանքներում:

6.1 Էյլերի հավասարումը

Դիցուք $L(x, y, y')$-ը՝ որպես երեք փոփոխականի ֆունկցիա, երկու անգամ անընդհատ դիֆերենցելի է R^3-ի վրա: Պահանջվում է գտնել այնպիսի $y(x) \in C^1[x_0, x_1]$ ֆունկցիա, որը բավարարի $y(x_0) = y_0$, $y(x_1) = y_1$ եզրային պայմաններին և հանդիսանա $I(y) \equiv \int_{x_0}^{x_1} L(x, y, y')\, dx$ ֆունկցիոնալի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ

$C^1[x_0, x_1]$ պարածության նորմի իմաստով: Այս խնդիրը ձևակերպվում է հետևյալ կերպ.

$$I(y) \equiv \int\limits_{x_0}^{x_1} L(x, y, y')\, dx \to extr, \qquad (6.1.1)$$

$$y(x_0) = y_0, \; y(x_1) = y_1,$$

որը կոչվում է ամրացված եզրերով վարիացիոն խնդիր: Այն վարիացիոն հաշվի պարզագույն խնդիրն է:

Այժմ տանք I ֆունկցիոնալի լոկալ մինիմումի (լոկալ մաքսիմումի) սահմանումը $C^1[x_0, x_1]$ պարածության նորմի իմաստով:

Սահմանում 6.1.1: *Դիցուք*

$$M \equiv \{y \in C^1[x_0, x_1]/y(x_0) = y_0, \; y(x_1) = y_1\}:$$

$y^* \in M$ կոչվում է I ֆունկցիոնալի լոկալ մինիմումի (լոկալ մաքսիմումի) կետ M բազմության վրա, եթե գոյություն ունի այնպիսի $\delta > 0$ թիվ, որ բոլոր $y \in$ $\in M$ ֆունկցիաների համար, որոնք բավարարում են $\|y - y^*\|_1 < \delta$ պայմանին, տեղի ունի

$$I(y) \geq I(y^*) \; (I(y) \leq I(y^*))$$

անհավասարությունը:
y^*-ը կոչվում է նաև (6.1.1) խնդրի լուծում:

Թեորեմ 6.1.1: *Եթե $y^*(x)$ ֆունկցիան (6.1.1) խնդրի լուծումն է, ապա այն բավարարում է*

$$-\frac{d}{dx} L'_{y'} + L'_y = 0 \qquad (6.1.2)$$

դիֆերենցիալ հավասարմանը, որը կոչվում է ***Էյլերի հավասարում:***

► Ենթադրենք, որ y^*-ը (6.1.1) խնդրում լոկալ մինիմումի կետ է (լոկալ մաքսիմումի դեպքը քննարկվում է անալոգ ձևով): Դիցուք $h(\cdot) \in C_0^1[x_0, x_1]$: Այսինքն՝

$$h(\cdot) \in C^1[x_0, x_1] \text{ և } h(x_0) = 0,\ h(x_1) = 0:$$

Դիտարկենք մեկ փոփոխականի հետևյալ ֆունկցիան.

$$\varphi(\alpha) = \int_{x_0}^{x_1} L(x, y^*(x) + \alpha h(x), (y^*(x))' + \alpha h'(x))\, dx: \quad (6.1.3)$$

Քանի որ y^*-ը I ֆունկցիոնալի լոկալ մինիմումի կետ է, ապա բավականաչափ փոքր α թվերի համար տեղի ունի

$$\varphi(\alpha) \geq \varphi(0)$$

անհավասարությունը: Այսինքն՝ 0 կետը φ ֆունկցիայի լոկալ մինիմումի կետ է:
Հետևաբար

$$\varphi'(0) = 0:$$

Այստեղից՝ ըստ պարամետրից կախված ինտեգրալի ածանցման կանոնի, կունենանք

$$\varphi'(0) = \int_{x_0}^{x^1} (L'_y h + L'_{y'} h')\, dx = 0: \quad (6.1.4)$$

Կատարենք մասերով ինտեգրում, հաշվի առնելով $h(x_0) = 0,\ h(x_1) = 0$ պայմանները՝ կստանանք

$$\int\limits_{x_0}^{x_1} L'_{y'} h' dx = L'_{y'}(h(x_1) - h(x_0)) - \int\limits_{x_0}^{x_1} \frac{d}{dx} L'_{y'} h\, dx =$$

$$= -\int\limits_{x_0}^{x_1} \frac{d}{dx} L'_{y'} h\, dx:$$

Այստեղից և (6.1.4)-ից՝ կստանանք

$$\int\limits_{x_0}^{x_1} (L'_y - \frac{d}{dx} L'_{y'}) h(x)\ dx = 0\ \forall h(x) \in C^1[x_0, x_1],$$

$$h(x_0) = 0,\ h(x_1) = 0: \tag{6.1.5}$$

Այժմ ցույց տանք, որ (6.1.5) պայմանից հետևում է Էյլերի հավասարումը:
Նշանակենք

$$a(x) \equiv L'_y(x, y^*, (y^*)') - \frac{d}{dx} L'_{y'}(x, y^*, (y^*)'):$$

Ցույց տանք, որ

$$a(x) = 0\ \forall x \in [x_0, x_1]:$$

Ենթադրենք, որ ինչ-որ $\xi \in [x_0, x_1]$ կետում $a(\xi) \neq 0$: Ընդհանրությունը չխախտելով՝ ենթադրենք, որ $a(\xi) > 0$: Քանի որ $a(x)$-ը անընդհատ ֆունկցիա է, ապա գոյություն կունենա ξ կետի այնպիսի շրջակայք՝ $(\xi - \delta, \xi + \delta) \subseteq [x_0, x_1]$, որ

$$a(x) > 0\ \forall x \in (\xi - \delta, \xi + \delta):$$

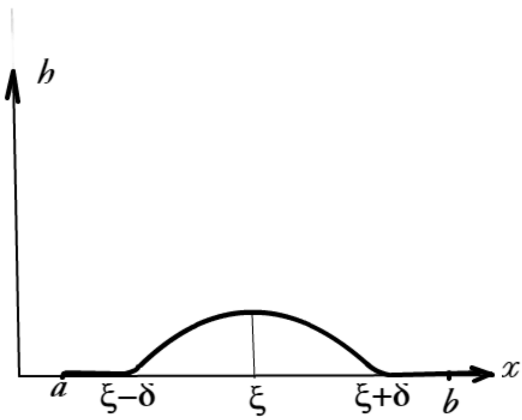
Դիցուք

$$p = \xi - \delta, \ q = \xi + \delta\text{:}$$

Այժմ դիտարկենք հետևյալ $h(x)$ ֆունկցիան.

$$h(x) = \begin{cases} 0, & \text{եթե } x \in [x_0, p] \\ (x-p)^2(x-q)^2, & \text{եթե } x \in [p, q] \\ 0, & \text{եթե } x \in [q, x_1]\text{:} \end{cases}$$

(տես գծ. 6.1)



Գծ. 6.1: h ֆունկցիայի գրաֆիկը

Պարզ է, որ $h(\cdot) \in C^1[x_0, x_1]$ և $h(x_0) = h(x_1) = 0$:

Ունենք
$$\int_{x_0}^{x_1} a(x)h(x)dx = \int_{p}^{q} a(x)h(x)\,dx > 0,$$
որը հակասում է (5.1.5)-ին: ■

Օրինակ (Էյլերի հավասարման լուծումը հանդիսանում է էքստրեմումի խնդրի լոկալ մինիմումի կետ):

$$I(y) \equiv \int_0^1 (y')^3\,dx \to min,\ y(0)=0,\ y(1)=1):$$

Կազմենք Էյլերի հավասարումը

$$\frac{d}{dx}3(y')^2 = 0 \Rightarrow 3(y')^2 = C \Rightarrow y' = const \Rightarrow$$
$$\Rightarrow y(x) = C_1x + C_2:$$

Հաշվի առնելով խնդրի եզրային պայմանները՝ կստանանք $y^*(x) = x$, որը խնդրի միակ էքստրեմալն է: Ցույց տանք, որ այն լոկալ մինիմումի կետ է:
Դիցուք
$$h(\cdot) \in C_0^1[0,1]:$$

Այդ դեպքում

$$I(y^*+h) = \int_0^1 (1+h')^3\,dx = \int_0^1 dx + 3\int_0^1 h'\,dx +$$
$$+\int_0^1 (h')^2(3+h')\,dx = I(y^*) + \int_0^1 (h')^2(3+h')\,dx:$$

Այստեղից՝ ակնհայտ է, որ եթե $\|h\|_1 < 3$, ապա

$$3 + h(x) > 0 \; \forall x \in [0,1]:$$

Հետևաբար

$$I(y^* + h) \geq I(y^*),$$

այսինքն՝ $y^*(\cdot)$-ը լոկալ մինիմումի կետ է:

ԽՆԴԻՐՆԵՐ

1. Լուծել վարիացիոն հաշվի հետևյալ պարզագույն խնդիրները.

ա) $\displaystyle\int\limits_0^1 ((y')^2 + y^2)\, dx \to min, \; y(0) = 0, \; y(1) = 1:$

բ) $\displaystyle\int\limits_{-1}^0 (12xy - (y')^2) dx \to min, \; y(-1) = 1, \; y(0) = 0:$

գ) $\displaystyle\int\limits_0^{\pi/2} ((y')^2 - y^2)\, dx \to min, \; y(0) = 1, \; y(\pi/2) = 0:$

դ) $\displaystyle\int\limits_0^1 (y^2 + (y')^2 + 2y\exp(x))\, dx \to min, \; y(0) = 0, \; y(1) = 1:$

ե) $\displaystyle\int\limits_0^{3/2} ((y')^2 + 2y)\, dx \to min, \; y(0) = 0, \; y(3/2) = 1:$

զ) $\int\limits_0^1 (4y\sin x - y^2 - (y')^2)\,dx \to max,\ y(0) = y(1) =$ $= 0$:

է) $\int\limits_0^{\pi/2} (6y\sin 2x + y^2 - (y')^2)\,dx \to max,\ y(0) =$ $= y(\pi/2) = 0$:

2. **Ամենաարագ վայրէջքի խնդիրը**: Ուղղաձիգ հարթության մեջ միևնույն ուղղաձիգի վրա չգտնվող $A(x_0, y_0)$ և $B(x_1, y_1)$ կետերը միացնել այնպիսի ողորկ կորով, որով ծանրության ուժի ազդեցությամբ շարժվող նյութական կետը վերևի A կետից առանց սկզբնական արագության կհասնի ներքևի B կետ ամենակարճ ժամանակում (տես գծ.6.2):

Լուծում: Դիցուք $y(x)$-ը ողորկ կոր է, որը միացնում է A և B կետերը: Դիցուք $M(x, y(x))$-ը կամայական կետ է կորի վրա: Ըստ էներգիայի պահպանման օրենքի՝ ունենք

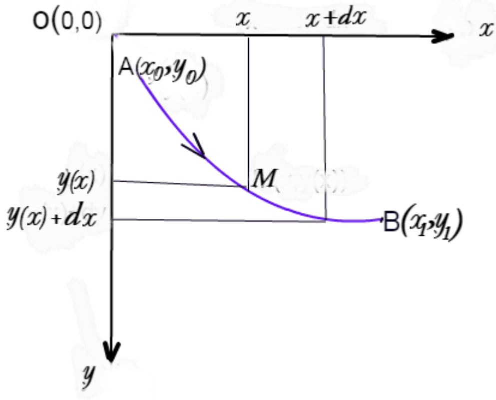
$$mv^2/2 = mgy(x),$$

որտեղ m-ը նյութական կետի մասսան է, իսկ v-ն՝ արագությունը M կետում, g-ն՝ ազատ անկման արագացումը: Այստեղից կստանանք

$$v = \sqrt{2gy(x)}:$$

Մյուս կողմից, ունենք

$$v = \frac{ds}{dt} = \frac{\sqrt{1 + (y')^2}dx}{dt},$$



Գծ. 6.2: Ամենաարագ վայրէջքի խնդիրը

որտեղ ds-ը էլեմենտար աղեղի երկարությունն է: Հետևաբար

$$T(y) = \frac{1}{\sqrt{2g}} \int_{x_0}^{x_1} \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{y(x)}} dx,$$

որտեղ $T(y)$-ը այն ժամանակամիջոցն է, որի ընթացքում կետը փրված կորով A կետից հասնում է B կետ: Քանի որ $1/\sqrt{2g} > 0$ հաստատուն է, ապա $T(y)$ ֆունկցիոնալի մինիմիզացիայի խնդրում կարելի է այն հաշվի չառնել:

Վերջնականորեն կստանանք էքստրեմումի հետևյալ խնդիրը.

$$I(y) \equiv \int\limits_{x_0}^{x_1} \frac{\sqrt{1+(y'(x))^2}}{\sqrt{y(x)}} dx \to min, \qquad (6.1.6)$$

$$y(x_0) = y_0, \ y(x_1) = y_1:$$

Կազմենք Էյլեի հավասարումը: Քանի որ (6.1.6)-ում ենթինտեգրալ L ֆունկցիոնալը բացահայտ կախված չէ x փոփոխականից, ապա Էյլերի հավասարումն ունի հետևյալ տեսքը.

$$L - y'L'_{y'} = C \text{ (տես, օրինակ՝ [2])}:$$

Այստեղից մեր օրինակի համար կունենանք

$$L - y'L'_{y'} = \frac{1+(y')^2}{\sqrt{y}} - y'\frac{y'}{\sqrt{1+(y')^2 y}} = C:$$

Պարզեցնելուց հետո կստանանք

$$y(1+(y')^2) = \frac{1}{C^2} = C_1:$$

Նշանակենք

$$y' = ctgt \Rightarrow y = \frac{C_1}{1+(ctgt)^2} = C_1 sin^2 t \Rightarrow$$

$$\Rightarrow dy = 2C_1 sintcostdt:$$

Այստեղից կստանանք

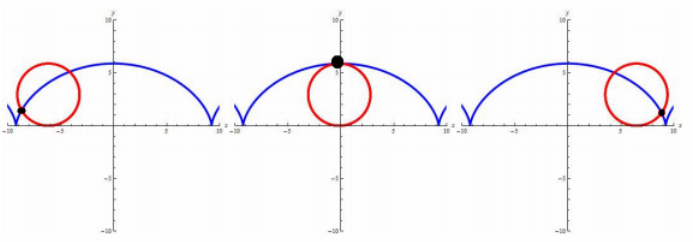
$$dx = \frac{dy}{y'} = \frac{2C_1 sintcostdt}{ctgt} = 2C_1 sin^2 tdt \Rightarrow$$

$$\Rightarrow x(t) = C_1/2(2t - sin2t) + C_2:$$

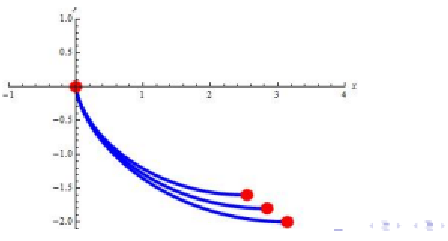
Նշանակելով $p = 2t$՝ կստանանք ցիկլոիդների ընտանիք

$$x = C_1/2(p - sinp) + C_2,\ y = C_1/2(1 - cosp):$$

C_1, C_2 հաստատունները որոշվում են այն պայմանից, որ ցիկլոիդը պետք է անցնի A և B կետերով: Ցիկլոիդը այն կորն է, որով շարժվում է շրջանագծի ֆիքսած կետը, երբ շրջանագիծը առանց սահքի գլորվում է հորիզոնական ուղղով (տես գծ.6.3): Կարելի է ցույց տալ, որ յուրաքանչյուր (x_1, y_1) $(x_1 > 0\ , y_1 > 0)$ կետի համար կա միայն մեկ ցիկլոիդ, որն է անցնում $A(x_0, y_0)$, $(y_0 > 0)$ և $B(x_1, y_1)$ կետերով (տես գծ.6.4):



Գծ. 6.3: Ցիկլոիդի կամար



Գծ. 6.4: Ամենաարագ վայրէջքի խնդրի լուծումը յուրաքանչյուր (x_1, y_1) կետի համար

3. Բոլոր ողորկ կորերի մեջ, որոնք միացնում են հարթության $A(2, 1)$ և $B(1, 0)$ կետերը, գտնել այն կորը, որով $v = x$

արագությամբ շարժվող նյութական կետը A կետից կհասնի B կետ ամենակարճ ժամանակում:

6.2 Լագրանժի մեթոդը վարիացիոն հաշվի խնդիրներում

Դիտարկենք հետևյալ խնդիրը.

$$I_0(y) \equiv \int\limits_{x_0}^{x_1} f_0(x, y, y')\, dx \to extr, \tag{6.2.1}$$

$$I_i(y) = \int\limits_{x_0}^{x_1} f_i(x, y, y')\, dx = \alpha_i,\ i \in [1 : m], \tag{6.2.2}$$

$$y(x_0) = y_0,\ y(x_1) = y_1: \tag{6.2.3}$$

Այս խնդիրը կոչվում է վարիացիոն հաշվի **իզոպերիմետրիկ**[1] խնդիր: Ենթադրվում է, որ $f_i,\ \ i \in [0 : m]$ ֆունկցիաները, որպես երեք փոփոխականի ֆունկցիաներ, երկու անգամ անընդհատ դիֆերենցելի են R^3-ի վրա, իսկ $\alpha_1, \alpha_2, ..., \alpha_m$ հաստատունները տրված թվեր են:
$y^{*1}[x_0, x_1]$ ֆունկցիան կոչվում է թույլատրելի, եթե այն բավարարում է (6.2. 2) և (6.2.3) պայմաններին:

Սահմանում 6.2.1: *Կասենք, որ թույլատրելի* y^* *ֆունկցիան (6.2.1)-(6.2.3) խնդրում լոկալ մինիմում է (լոկալ մաքսիմում է), եթե գոյություն ունի այնպիսի* $\delta > 0$ *թիվ, որ բոլոր թույլատրելի* y *ֆունկցիաների համար,*

[1] միևնույն պարամետր ունեցող

որոնք բավարարում են $\|y-y^*\|_1<\delta$ *պայմանին, տեղի ունի*

$$I_0(y)\geq I_0(y^*)\ (I_0(y)\leq I_0(y^*))$$

անհավասարությունը:
Կազմենք Լագրանժի ֆունկցիան

$$L(x,y,y',\lambda)\equiv\lambda_0 f_0(x,y,y')+\lambda_1 f_1(x,y,y')+...+$$

$$+\lambda_m f_m(x,y,y'),$$

որտեղ $\lambda=(\lambda_0,\lambda_1,...,\lambda_m)\in R^{m+1}$:

Թեորեմ 6.2.1: *Դիցուք* $y^*(x)$ *ֆունկցիան (6.2.1) խնդրի լուծում է: Այդ դեպքում գոյություն ունեն այնպիսի* $\lambda_0,...,\ \lambda_m$ *թվեր, որոնցից գոնե մեկը զրո չէ, որ* $y^*(x)$*-ը բավարարում է*

$$-\frac{d}{dx}L'_{y'}+L'_y=0$$

դիֆերենցիալ հավասարմանը $y(x_0)=y_0,\ \ y(x_1)=y_1$ եզրային պայմաններով:
► Դիցուք y^* ֆունկցիան (6.2.1)-(6.2.3) խնդրում լոկալ մինիմում է (լոկալ մաքսիմումի դեպքը քննարկվում է անալոգ ձևով): Նշանակենք

$$\delta I_i(y^*,h)\equiv\lim_{\alpha\downarrow 0}\frac{I_i(y^*+\alpha h)-I_i(y^*)}{\alpha}:$$

Հեշտ է ցույց տալ, որ

$$\delta I_i(y^*,h)=\int\limits_{x_0}^{x_1}(-\frac{d}{dx}f'_{iy'}+f'_{iy})h(x)\,dx,\ \ i\in[0:m]:$$

Այսպեղից հեպևում է, որ $\delta I_i(y^*, h)$ ֆունկցիոնալը h-ի նկապմամբ գծային է: Դիպարկենք հեպևյալ գծային օպերապորը.

$$A : C_0^1 \to R^{m+1},$$

$$Ah \equiv (\delta I_0(y^*, h), \delta I_1(y^*, h), ..., \ \delta I_m(y^*, h)):$$

ImA-ով նշանակենք A օպերապորի պապկերը: Հնարավոր է երկու դեպք.

1) $ImA \subset R^{m+1}$,

2) $ImA = R^{m+1}$:

Առաջին դեպքում ImA-ն R^{m+1} պարածության սեփական ենթապարածությունն է: Հեպևաբար, գոյություն ունի ոչ զրոյական այնպիսի $\lambda = (\lambda_0, ..., \ \lambda_{m+1})$ վեկպոր, որը ուղղահայաց է այդ ենթապարածությանը, այսինքն՝

$$(\lambda, Ah) = 0 \ \Rightarrow \lambda_0 \delta I_0(y^*, h) + ... \ + \lambda_m \delta I_m(y^*, h) = 0$$

$$\forall h \in C_0^1:$$

Այսպեղից կսպանանք

$$\int_{x_0}^{x_1} (-\frac{d}{dx} L'_{y'} + L'y) h(x) \, dx = 0 \ \forall h \in C_0^1: \qquad (6.2.2)$$

Ինչպես պարզագույն խնդրում, այսպեղ նույնպես կարելի է ցույց պալ, որ (6.2.2)-ից հեպևում է, որ

$$-\frac{d}{dx} L'_{y'} + L'y = 0:$$

Դիտարկենք երկրորդ դեպքը: Դիցուք $\{e_0, ..., \quad e_m\}$ համակարգը բազիս է կազմում R^{m+1} տարածությունում: Ընտրենք այնպիսի $h^0, h_1, ..., \ h_m$ ֆունկցիաներ, որ

$$\delta I_j(y^*, h_i) = \begin{cases} 1, & \text{եթե } i = j, \\ 0, & \text{եթե } i \neq j: \end{cases}$$

Կազմենք հավասարումների հետևյալ համակարգը.

$$\begin{cases} \varphi_0(\beta_0, ..., \ \beta_m) \equiv I_0(y^* + \sum_{j=0}^m \beta_j h_j) = I_0(y^*) - \varepsilon, \\ \\ \varphi_i(\beta_0, ..., \ \beta_m) \equiv I_i(y^* + \sum_{j=0}^m \beta_j h_j) = \alpha_i, \ i \in [1 : m]: \end{cases} \tag{6.2.3}$$

Այս համակարգում ε-ը պարամետր է: Ունենք.

$$\frac{\partial \varphi_i}{\partial \beta_j}(0) = \delta I_i(y^*, h_j):$$

Այստեղից հետևում է, որ (6.2.3) համակարգը բավարարում է հակադարձ արտապատկերումների մասին թեորեմի բոլոր պայմաններին (տե՛ս, օրինակ՝ [3], էջ 31): Դա նշանակում է, որ գոյություն ունի այնպիսի $\beta(\varepsilon) \equiv (\beta_0(\varepsilon), ..., \ \beta_m(\varepsilon))$ վեկտոր ֆունկցիա, որոշված զրոյի ինչ-որ մի շրջակայքում, որը այդտեղ բավարարում է (6.2.3) համակարգին և $\beta(\varepsilon) \to$ $\to 0$, երբ $\beta \to 0$: Հետևաբար, եթե ε պարամետրը դրական է, ապա կստանանք հակասություն, քանի որ y^*-ը (6.2.1) խնդրի լոկալ մինիմումի կետ է:
■

Օրինակ (Էյլերի հավասարման լուծումը իզոպերիմետրիկ խնդրում գլոբալ մինիմումի կետ է):

$$I_0(y) = \int_0^1 (y')^2 \to min,$$

$$I_1(y) = \int_0^1 y\,dx = 0,\ y(0) = 0,\ y(1) = 1:$$

Լուծում: Կազմենք Լագրանժի ֆունկցիան

$$L = \lambda_0 (y')^2 + \lambda_1 y:$$

Էյլերի հավասարումը այս ֆունկցիայի համար հետևյալն է.

$$-\frac{d}{dx} L'_{y'} + L'_y = 0 \Rightarrow -2\lambda_0 y'' + \lambda_1 = 0:$$

Եթե $\lambda_0 = 0$, ապա $\lambda_1 = 0$ և Լագրանժի բոլոր գործակիցները հավասար են զրոյի: Այդ դեպքում թույլատրելի էքստրեմալներ չկան: Էյլերի հավասարման մեջ տեղադրենք $\lambda_0 = 1$: Այդ դեպքում այդ հավասարման ընդհանուր լուծումը կլինի $y(x) = C_1 x^2 + C_2 x + C_3$ ֆունկցիան: C_1, C_2, C_3 անորոշ գործակիցները որոշենք եզրային և իզոպերիմետրիայի հետևյալ պայմաններից.

$$\begin{cases} y(0) = 0 \Rightarrow C_3 = 0, \\ y(1) = 1 \Rightarrow C_1 + C_2 = 1, \\ \int_0^1 y dx = 0 \Rightarrow \int_0^1 (C_1 x^2 + C_2 x)\, dx = 0 \Rightarrow \end{cases}$$

$$\Rightarrow C_1/3 + C_2/2 = 0:$$

Այստեղից կստանանք միակ թույլատրելի էքստրեմալը՝

$$y^*(x) = 3x^2 - 2x:$$

Ապացուցենք, որ այն գլոբալ մինիմումի կետ է: Դիցուք $y(\cdot)$-ը թույլատրելի ֆունկցիա է: Այդ դեպքում

$$y(\cdot) - y^*(\cdot) = h(\cdot) \in C_0^1[0,1] \text{ և } \int_0^1 h\,dx = 0:$$

Ունենք՝

$$I_0(y(\cdot)) - I_0(y^*(\cdot)) = \int\limits_0^1 ((y^*)' + h')^2\, dx - \int\limits_0^1 ((y^*)')^2\, dx =$$

$$= \int\limits_0^1 2(y^*)'h dx + \int\limits_0^1 (h')^2 dx \geq 2\int\limits_0^1 (y^*)'h'\, dx:$$

Կապարելով մասերով ինտեգրում՝ կստանանք

$$\int\limits_0^1 (y^*)'h'\, dx = \int\limits_0^1 (y^*)\, dh = y^*h|_0^1 - \int\limits_0^1 (y^*)''h\, dx =$$

$$= -6\int\limits_0^1 h\, dx = 0:$$

Այսպիսով,

$$I_0(y^*(\cdot)) \geq I_0(y^*(\cdot))$$

ցանկացած թույլատրելի $y(\cdot)$ ֆունկցիայի համար:

Օրինակ (Դիդոնայի[2] խնդիրը մաքսիմալ մակերեսով սեղա– նակերպի մասին):

Տրված է $f(x)$ ֆունկցիան $[-x_0, x_0]$ հատվածի վրա: Գրաֆիկը ներկայացնող կորի երկարությունը հաստատուն է: Գտնել գրաֆիկի տեսքը այնպես, որ կորագիծ սեղանի մակերեսը լինի մեծագույն:

[2] Կարթագենի թագուհի, մ.թ.ա. 8-րդ դար

Այս խնդիրը ձևակերպվում է հետևյալ կերպ.

$$\int_{-x_0}^{x_0} y(x)\,dx \to max,$$

$$\int_{-x_0}^{x_0} \sqrt{1+(y')^2}\,dx = l, y(-x_0) = y(x_0) = 0:$$

Լուծում: Կազմենք Լագրանժի ֆունկցիան.

$$L(y, y', \lambda) = \lambda_0 y + \lambda_1 \sqrt{1+(y')^2}:$$

Քանի որ Լագրանժի ֆունկցիան բացահայտ կախված չէ x փոփոխականից, ապա Էյլերի հավասարումը կունենա հետևյալ տեսքը.

$$L - y' L'_{y'} = C \Rightarrow \lambda_0 y - C = \frac{-\lambda_1}{\sqrt{1+(y')^2}}: \qquad (6.2.4)$$

Այստեղից հետևում է, որ եթե $\lambda_0 = 0$, ապա կամ $\lambda_1 = 0$ կամ $y' = 0$:
Առաջին դեպքը հնարավոր չէ, որովհետև Լագրանժի գործակիցները միաժամանակ զրո լինել չեն կարող: Երկրորդ դեպքում, հաշվի առնելով եզրային և իզոպերիմետրային պայմանները, կունենանք

$$y^*(x) \equiv 0,\ l = 2x_0:$$

Եթե $\lambda_0 = 1$, ապա նշանակելով $y' = tgt$, (6.2.4)-ից կստանանք՝

$$y(t) - C = -\lambda_1 \cos t: \qquad (6.2.5)$$

Մյուս կողմից, ունենք

$$\frac{dy}{dx} = tgt \Rightarrow dx = \frac{dy}{tgt} \Rightarrow x(t) - C_1 = \lambda_1 \sin t: \qquad (6.2.6)$$

(6.2.5)-(6.2.6) հավասարություններից հետևում է, որ

$$(x - C_1)^2 + (y - C)^2 = \lambda_1^2:$$

Եզրային պայմաններից ստանում ենք, որ $C_1 = 0$: Այսպիսով, եթե $l < 2x_0$, ապա խնդիրը լուծում չունի: Եթե $l = 2x_0$, ապա $y^*(x) \equiv 0$: Եթե $l > 2x_0$ և խնդիրը ունի օպտիմալ լուծում, ապա նրա գրաֆիկը պետք է ունենա շրջանագծային աղեղի տեսք: Այդ շրջանագիծը անցնում է $(-x_0, 0)$ և $(x_0, 0)$ կետերով, իսկ նրա կենտրոնը գտնվում է OY առանցքի վրա: Կարելի է ցույց տալ, որ եթե $l > \pi x_0$, ապա խնդիրը օպտիմալ լուծում չունի:

ԽՆԴԻՐՆԵՐ

Լագրանժի գործակիցների մեթոդով լուծել վարիացիոն հաշվի հետևյալ իզոպերիմետրիկ խնդիրները:

ա) $\int\limits_0^1 (y')^2\,dx \to min, \ \int\limits_0^1 y\,dx = 0, \ y(0) = 0, \ y(1) = 1:$

բ) $\int\limits_0^1 (y')^2\,dx \to min, \ \int\limits_0^1 xydx = 0, \ y(0) = y(1) = 0:$

գ) $\int\limits_0^\pi (y')^2\,dx \;\to\; min,\;\; \int\limits_0^\pi ysinx\,dx \;=\; 0,\;\; y(0) \;=$
$= y(\pi) = 1$:

դ) $\int\limits_0^\pi (y')^2\,dx \;\to\; min,\;\; \int\limits_0^\pi ycosxdx \;=\; \pi/2,\;\; y(0) \;=$
$= 1,\; y(\pi) = -1$:

ե) $\int\limits_0^\pi ysinx\,dx \;\to\; min,\;\; \int\limits_0^\pi (y^{'})^2\,dx \;=\; 3\pi/2,\;\; y(0) \;=$
$= 1,\; y(\pi) = \pi$:

6.3 Վարիացիոն հաշվի դասական իզոպերիմետրիկ խնդիրը

Խնդիր: Ապացուցել, որ l երկարության կտոր առ կտոր ողորկ, պարզ հարթ փակ կորերի մեջ ամենամեծ մակերես զբաղեցնում է շրջանագիծը:

Լուծում: Դիցուք

$$x = x(s),\;\; y = y(s),\; s \in [0, l]$$

կորի պարամետրական հավասարումներն են: Հաշվի առնելով փոփոխական աղեղի երկարության դիֆերենցիալի և մակերեսի հայտնի բանաձևերը՝ կարելի է տալ խնդրի հետևյալ մաթեմատիկական ձևակերպումը.

$$S(x, y) = \int\limits_0^l x(s)\frac{dy}{ds}\,ds \to max,\;\; (\frac{dx}{ds})^2 + (\frac{dy}{ds})^2 = 1:$$

$x(s)$ և $y(s)$ ֆունկցիաները ներկայացնենք Ֆուրիեի շարքով՝

$$x(s) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos\frac{2\pi n}{l}s + b_n \sin\frac{2\pi n}{l}s), \qquad (6.3.1)$$

$$y(s) = \frac{1}{2}c_0 + \sum_{n=1}^{\infty}(c_n \cos\frac{2\pi n}{l}s + d_n \sin\frac{2\pi n}{l}s))։ \qquad (6.3.2$$

Այստեղից՝ այս ֆունկցիաների ածանցիալների համար կունենանք հետևյալ բանաձևերը.

$$\frac{dx}{ds} = \sum_{n=1}^{\infty}(-\frac{2\pi n}{l}a_n \sin\frac{2\pi n}{l}s + \frac{2\pi n}{l}b_n \cos\frac{2\pi n}{l}s), \qquad (6.3.3)$$

$$\frac{dy}{ds} = \sum_{n=1}^{\infty}(-\frac{2\pi n}{l}c_n \sin\frac{2\pi n}{l}s + \frac{2\pi n}{l}d_n \cos\frac{2\pi n}{l}s։ \qquad (6.3.4)$$

Հայտնի է նաև, որ եթե $\alpha_n,\ \beta_n,\ n = 0, 1, 2, ...$ թվերը f ֆունկցիայի Ֆուրիեի գործակիցներն են, իսկ $\gamma_n,\ \delta_n,\ n = 0, 1, 2, ...$ թվերը՝ φ-ի գործակիցներն են, ապա

$$\frac{2}{l}\int_0^l f(s)ds = \frac{\alpha_0^2}{2} + \sum_{i=1}^{\infty}(\alpha_n^2 + \beta_n^2),$$

$$\frac{2}{l}\int_0^l f(s)\varphi(s)ds = \frac{1}{2}\alpha_0\gamma_0 + \sum_{n=1}^{\infty}(\alpha_n\gamma_n + \beta_n\delta_n)։$$

Այստեղից՝ նկատի ունենալով նաև (6.3.3) և (6.3.4) բանաձևերը՝ հարթ պատկերի մակերեսի համար կստանանք հետևյալ բանաձևը.

$$S = \pi\sum_{n=1}^{\infty} n(a_n d_n - b_n c_n)\ ։ \qquad (6.3.5)$$

Քանի որ

$$\int\limits_0^l ((\frac{dx}{ds})^2 + (\frac{dy}{ds})^2)\, ds = l,$$

ապա, հաշվի առնելով նաև ածանցիալների (6.3.3)-(6.3.5) բանաձևերը, կստանանք, որ կորի l երկարությունը պետք է բավարարի հետևյալ հավասարմանը.

$$l = \frac{2\pi^2}{l}\sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2 + c_n^2 + d_n^2)\text{:} \qquad (6.3.6)$$

Մակերեսի (6.3.5) և կորի երկարության (6.3.6) բանաձևերից հետևում է, որ

$$\frac{l^2}{4\pi} - S = \frac{\pi}{2}\sum_{n=1}^{\infty}((na_n - d_n)^2 + (nb_n + c_n)^2 +$$

$$+(n^2-1)(c_n^2 + d_n^2)) \geq 0\text{:} \qquad (6.3.7)$$

Այսպիսով, ստացանք հանհրահայտ հետևյալ իզոպերիմետրիկ անհավասարությունը.

$$S \leq \frac{l^2}{4\pi}\ \text{:} \qquad (6.3.8)$$

Պարզ է, որ (6.3.7) անհավասարությունը կվերածվի հավասարության, երբ

$$a_1 = d_1,\ b_1 + c_1 = 0,\ a_n = b_n = c_n = d_n = 0,$$

$$n = 2, 3, \ldots\text{:}$$

Այստեղից՝

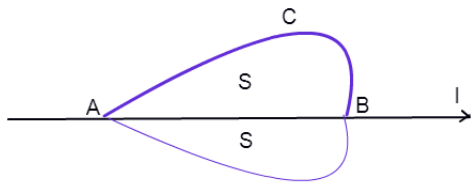
$$x = \frac{1}{2}a_0 + a_1\cos s + b_1 \sin s,$$

$$y = \frac{1}{2}c_0 - b_1 \cos s + a_1 \sin s,$$

այսինքն՝

$$(x - \frac{1}{2}a_0)^2 + (y - \frac{1}{2}c_0)^2 = a_1^2 + b_1^2 = \frac{l^2}{4\pi}:$$

Հետևանք 6.3.1 *(Դիդոնայի առաջին խնդիրը): L երկարություն ունեցող բոլոր ուղղելի կորերի մեջ գտնել այն կորը, որի ծայրերը գտնվում են l ուղղի վրա և, որը $[A, B]$ հատվածի հետ վերին կիսահարթությունում սահմանափակում է ամենամեծ մակերեսով պատկեր (տես գծ.6.5) :*



Գծ. 6.5: Դիդոնայի առաջին խնդիրը

Դիցուք ACB-ն կամայական աղեղ է A և B սկզբնակետերով և $[A, B]$ հատվածի հետ սահմանափակում է S մակերես (տես. գծ.6.5): ACB աղեղը համաչափ արտապատկերելով l ուղղի նկատմամբ՝ ստանում ենք $2L$ երկարությամբ փակ կոր, որը սահմանափակում է $2S$ մակերեսով պատկեր: Համաձայն արդեն հայտնի (6.3.8) իզոպերիմետրիկ անհավասարության՝ կստանանք

$$(2L)^2 \geq 4\pi 2S:$$

Որտեղից

$$S \leq \frac{L^2}{2\pi}:$$

Հետևաբար, S-ի մաքսիմալ արժեքը հավասար կլինի $\frac{L^2}{2\pi}$ և այն հասանելի կլինի, երբ ACB աղեղը լինի $[AB]$ տրամագծով կիսաշրջան:

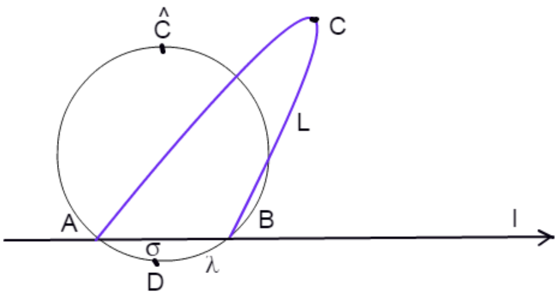
Հետևանք 6.3.2 *(Դիդոնայի երկրորդ խնդիրը): Դիդոնայի առաջին խնդրում աղեղի A, B ծայրակետերը կամայական են l ուղղի վրա (նրանք շարժական են): Այժմ դիտարկենք նույն խնդիրը միայն այն ենթադրությամբ, որ աղեղի A, B ծայրակետերը ամրագրված են l ուղղի վրա:*

Դիցուք $A\hat{C}B$ աղեղը շրջանագծային աղեղ է հենված $[A, B]$ լարի վրա (տե՛ս գծ.6.6): Լրացնենք այն ADB աղեղով մինչև շրջանագիծ: ADB աղեղի երկարությունը նշանակենք λ-ով, իսկ այդ աղեղով և $[A, B]$ հատվածով սահմանափակված սեգմենտի մակերեսը՝ σ-ով: Դիցուք ACB-ն կամայական աղեղ է՝ բավարարող խնդրի պայմաններին և $[A, B]$ հատվածի հետ եզրափակում է S մակերեսով պատկեր: $ACBD$ փակ կորը ունի $L + \lambda$ երկարություն և սահմանափակում է $S + \sigma$ մակերես: Ըստ իզոմետրիկ (6.3.8) անհավասարության՝ ունենք

$$4\pi(S + \sigma) \leq (L + \lambda)^2:$$

Այստեղից՝

$$S \leq \frac{1}{4\pi}(L + \lambda)^2 - \sigma:$$



Գծ. 6.6: Դիդոնայի երկրորդ խնդիրը

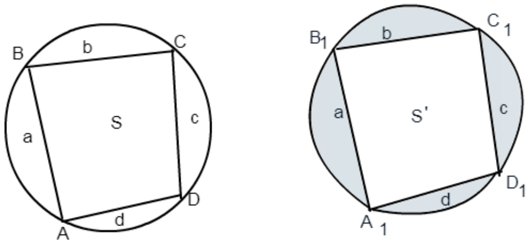
Հետևաբար S-ը կհասնի մաքսիմումի, երբ $ACBD$ կորը լինի շրջանագիծ:

ԽՆԴԻՐՆԵՐ

---

1. Դիցուք ունենք երկու քառանկյուններ, որոնց կողերի երկարությունները միևնույն a, b, c, d թվերն են: Ենթադրենք նրանցից մեկին կարելի է արտագծել շրջանագիծ: Ապացուցել, որ նրա մակերեսը մեծ է կամ հավասար մյուս քառանկյան մակերեսից:

   Ցուցում: $A_1B_1C_1D_1$ քառանկյան կողերին «կպցնել» $ABCD$ քառանկյան համապատասխան



Գծ. 6.7:

կողերին կից սեգմենտները և օգտվել իզոպերի–մետրիկ անհավասարությունից (տես գծ.6.7):

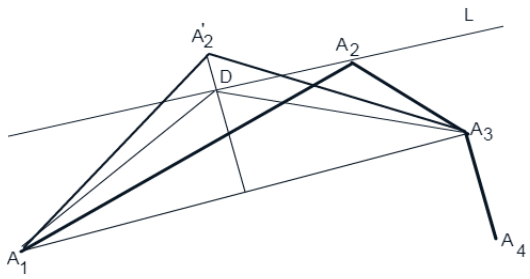
2. Ցույց տալ, որ շրջանին ներգծված n անկյուն բազմանկյունների մեջ ամենամեծ մակերես ունի կանոնավորը:

Ցուցում: Եթե $A_1A_2 \neq A_2A_3$, ապա L $(L \parallel A_1A_3)$ ուղղի վրա ընտրել D կետը

$$\mid A_1D \mid + \mid DA_3 \mid \to min$$

պայմանից: Այնուհետև, ցույց տալ, որ CD ճառագայթի վրա կարելի է վերցնել A_2' կետը այնպես, որ

$$S_{A_1A_2'A_3} > S_{A_1A_2A_3} \text{ (տես գծ.6.8)}:$$



Գծ. 6.8:

3*. Տրված պարագծով ուռուցիկ n անկյուն բազմ– անկյունների մեջ գտնել այն բազմանկյունը, որի մակերեսն ամենամեծն է (Զենոդորի խնդիր):

4. Ցույց տալ, որ միևնույն հիմքով և տրված պարագիծն ունեցող եռանկյունների մեջ ամենամեծ մակերեսն ունի հավասարասրուն եռանկյունին:

6.4 Էքստրեմումի բավարար պայմանները վարիացիոն հաշվի խնդիրներում

Այժմ բերենք վարիացիոն հաշվի պարզագույն խնդրի օրինակ, որտեղ Էյլերի հավասարումը ունի միակ լուծում, որը էքստրեմում չէ: Այսինքն՝ Էյլերի հավասարումը էքստրեմումի միայն անհրաժեշտ պայման է: Դիտարկենք հետևյալ խնդիրը.

$$I(y) \equiv \int\limits_0^{3\pi/2} ((y')^2 - y^2)\, dx \to min,\ y(0) = y(3\pi/2) = 0:$$

Էյլերի հավասարումը ունի հետևյալ տեսքը.

$$y'' + y = 0 \Rightarrow y(x) = C_1 \sin x + C_2 \cos x:$$

Հաշվի առնելով խնդրի եզրային պայմանները՝ կստանանք $y^*(x) \equiv 0$: Դիտարկենք ֆունկցիաների

$$y_n(x) = \frac{1}{n} \sin(\frac{2x}{3})$$

հաջորդականությունը: Ակնհայտ է, որ այս ֆունկցիաները թույլատրելի են և

$$y_n(\cdot) \xrightarrow{C^1} y^*(\cdot) :$$

Մյուս կողմից ունենք

$$I(y_n) = -\frac{5\pi}{n^2} < 0 = I(y^*),$$

այսինքն՝ y^*-ը լոկալ մինիմումի կետ չէ:
Այժմ բերենք բավարար մի պայման, որով սպուզվում է, թե երբ Էյլերի հավասարման լուծումը կլինի լոկալ մինիմումի կետ: Դիցուք $y^*(x)$-ը վարիացիոն հաշվի պարզագույն խնդրի Էյլերի հավասարման լուծում է: Նշանակենք

$$P(x) \equiv L''_{y'y'}(x, y^*(x), (y^*(x))'),$$

$$Q(x) = -\frac{d}{dx} L''_{yy'}(x, y^*(x), (y^*(x))')+$$

$$+L''_{yy}(x, y^*(x), (y^*(x))'):$$

Կազմենք հետևյալ դիֆերենցիալ հավասարումը, որը կոչվում է Յակոբիի հավասարում.

$$\frac{d}{dx}(P\frac{dh}{dx}) - Qh = 0 : \qquad (6.4.1)$$

Դիցուք $y^*(x)$ էքստրեմալը բավարարում է հետևյալ երկու պայմաններին.

- $y^*(x)$-ը բավարարում է **Յակոբիի պայմանին**, եթե (6.4.1) հավասարումը սկզբնական $h(x_0) = 0$, $h'(x_0) = 1$ պայմաններով ունի ոչ փրիվիալ այնպիսի $h(x)$ լուծում, որ

$$h(x) \neq 0 \;\; \forall\, x \in (x_0, x_1]:$$

- $y^*(x)$-ը բավարարում է **Լեժանդրի պայմանին**, եթե

$$P(x) > 0 \; \forall x \in [x_0, x_1]:$$

Ճշմարիտ է հետևյալ պնդումը (տես, օրինակ՝ [2-3]):

Թեորեմ 6.4.1: *Դիցուք $y^*(x)$-ը (6.1.1) վարիացիոն հաշվի պարզագույն խնդրի Էյլերի հավասարման լուծումն է և բավարարվում են Յակոբիի և Լեժանդրի պայմանները:*
Այդ դեպքում $y^*(x)$-ը (6.1.1) խնդրում լոկալ մինիմումի կետ է:

Օրինակ (Կարճագույն ճանապարհի խնդիրը):

$$\int_0^1 \sqrt{1 + (y')^2}\, dx \to min, \; y(0) = 0, \; y(1) = 1: \qquad (6.4.2)$$

Լուծում: Քանի որ այստեղ L ֆունկցիան բացահայտ կախված չէ x փոփոխականից, ապա Էյլերի հավասարումը ունի հետևյալ տեսքը.

$$L - y'L'_{y'} = C \Rightarrow \sqrt{1+(y')^2} - \frac{(y')^2}{\sqrt{1+(y')^2}} = C \Rightarrow$$

$$\Rightarrow y(x) = C_1 x + C_2 \text{:}$$

Հաշվի առնելով (6.4.2) եզրային պայմանները՝ կստանանք

$$y^*(x) = x \text{:}$$

Ստուգենք Յակոբիի պայմանը:
Ունենք

$$Q(x) = L''_{yy} - \frac{d}{dx} L''_{yy'} \equiv 0,$$

$$P(x) = \frac{1}{(1+(y^*)')^{3/2}} = \frac{1}{2\sqrt{2}} \text{:}$$

Յակոբիի հավասարումը ունի հետևյալ տեսքը.

$$\frac{d^2h}{dx^2} = 0 \text{:}$$

Այստեղից հետևում է, որ

$$h(x) = C_1 x + C_2 \text{:}$$

Ուստի $h(x) = x$ ֆունկցիան $h(0) = 0,\ h'(0) = 1$ սկզբնական պայմանով Յակոբիի հավասարման ոչ տրիվիալ լուծում է, որը զրո չի դառնում $(0, 1]$ կիսաբաց միջակայքի ոչ մի կետում: Հետևաբար Յակոբիի բավարար պայմանը տեղի ունի: Լեժանդրի պայմանը նույնպես տեղի ունի, քանի որ $P(x) \equiv 1/(2\sqrt{2}) > 0$: Այսպիսով, $y^*(x) = x$ ֆունկցիան կարճագույն ճանապարհի խնդրի լուծումն է:

ԽՆԴԻՐՆԵՐ

Օգտագործելով Յակոբիի և Լեժանդրի պայմանները՝ լուծել վարիացիոն հաշվի հետևյալ խնդիրները:

ա) $\int\limits_0^{\pi/2} ((y')^2 - y^2)\, dx \to min,\ y(0) = 0,\ y(\pi/2) = 1$:

բ) $\int\limits_0^{\pi/2} ((y')^2 - y^2 + 4ycosx)\, dx \to min,\ y(0) = 0,\ y(\pi/2) = \pi/2$:

գ) $\int_0^1 ((y')^2 + y^2 + 4ysh(x))dx \to min,\ y(0) = -1,\ y(1) = 0$:

դ) $\int\limits_0^{\pi/2} ((y')^2 - 4y^2)\, dx \to min,\ y(0) = 0,\ y(\pi/4) = 1$:

ե) $\int\limits_0^{\pi/2} (y^2 - 2(y')^2 + 2y)\, dx \to min,\ y(0) = y(\pi/2) = 0$:

Գլուխ 7

Կառավարում և օպտիմալ կառավարում գծային համակարգերում

Օպտիմալ կառավարման խնդիրները օպտիմիզացիայի տեսության խնդիրների լայն դաս է և ունեն կիրա–ռական կարևոր նշանակություն: Օպտիմալ կառավարման տեսությունը հենված է Պոնտրյագինի մաքսիմումի սկզբունքի վրա և ինտենսիվ զարգանում է: Հարկ ենք համարում նշել, որ օպտիմալ կառավարման տեսության հիմնադիրներն են Լ.Ս. Պոնտրյագինն ու նրա դպրոցը (Բոլտյանսկի, Միշենկո, Գամկրելիձե և ուրիշներ, 1953-1961, տես [19]), իսկ այդ տեսության տարբեր ասպեկտների և դրանց վերաբերյալ լայն գրականության հետ կարելի է ծանոթանալ [3, 4, 15, 31] աշխատանքներում:

Այս գլխում մենք կծանոթանանք կառավարման և օպտիմալ արագագործության խնդրի հետ գծային համակարգերում: Հենման ֆունկցիաների օգնությամբ այդ դասի խնդիրների համար կապացուցենք Պոնտրիագինի մաքսիմումի սկզբ–

բունքը որպես էքստրեմումի անհրաժեշտ և բավարար պայման:

7.1 Կառավարման և օպտիմալ կառավարման խնդիրների դրվածքները

Դիտարկելու ենք օբյեկտ, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով՝

$$\frac{dx}{dt} = f(t, x, u),$$

որտեղ u-ն պարամետր է և կոչվում է կառավարում: Կառավարվող օբյեկտի կոնկրետ շարժումը նկարագրելու համար անհրաժեշտ է

- տալ $u = u(t)$ կառավարումը որպես ֆունկցիա t ժամանակից:
- Նշել օբյեկտի սկզբնական վիճակը՝ $x(t_0)$-ն:
- Լուծել Կոշիի խնդիրը.

$$\frac{dx}{dt} = f(t, x(t), u(t)),\ x(t_0) = x^0: \qquad (7.1.1)$$

Ակնհայտ է, որ շարժման $x(t)$ հետագիծը կախված է $u(t)$ կառավարումից:

Ենթադրվում է, որ կառավարման $u(t)$ վեկտորը յուրաքանչյուր t պահի բավարարում է $u(t) \in U$ պայմանին, որտեղ $U \subset R^m$ կոմպակտ բազմություն է: $u(t)$ կառավարումը վերցնում են ֆունկցիաների տարբեր դասերից՝ դիֆերենցելի, անընդհատ, կտոր առ կտոր անընդհատ, չափելի և այլն: Ենթադրվում է նաև, որ կա սկզբնական մի M_0 բազմություն, որից սկսվում է օբյեկտի շարժումը, այսինքն՝ շարժման $x(t)$ հետագիծը սկզբնական t_0 պահին պետք է բավարարի $x(t_0) \in M_0$ պայմանին: Այստեղ դրվում է կառավարման և օպտիմալ կառավարման հետևյալ խնդիրները:

- **Անհրաժեշտ է** օբյեկտը սկզբնական M_0 բազմությունից թույլատրելի $u(t) \in U$ կառավարման օգնությամբ տեղափոխել նպատակային $M_1 \subset R^n$ բազմություն: Այսինքն՝ պետք է ընտրել այնպիսի թույլատրելի $u(t)$ կառավարում, որ այդ կառավարմանը համապատասխանող կոշու խնդրի լուծման $x(t)$ հետագիծը որևէ t_1 պահի հատի M_1 բազմությունը՝

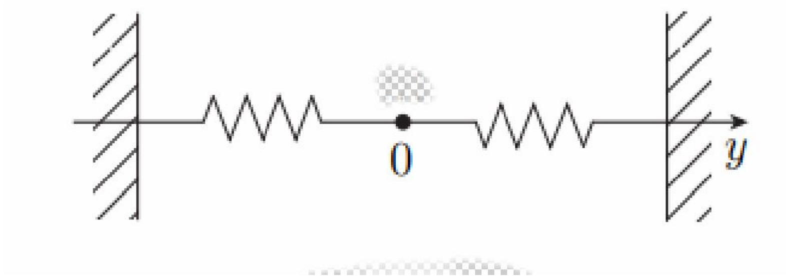
$$x(t_1) \in M_1:$$

  Սա կառավարման խնդրի ընդհանուր դրվածքն է:

- Օպտիմալ կառավարման **արագագործության** խնդիրը հետևյալն է. պետք է օբյեկտը M_0 բազմությունից տեղափոխել վերջնական M_1 բազմություն մինիմալ ժամանակամիջոցում: Այսինքն՝ $u(t)$ կառավարումը ընտրել այնպես, որ

$$x(t_1) \in M_1,\ t_1 - t_0 \to min:$$

Օրինակ: Դիտարկենք ֆիզիկական ճոճանակ, որը գտնվում է հավասարակշռության վիճակում: Եթե t պահին $y(t)$-ն ճոճանակի շեղումն է հավասարակշռության վիճակից, ապա նրա շարժման հավասարումը նկարագրվում է հետևյալ



Գծ. 7.1:

դիֆերենցիալ հավասարումով.

$$\frac{d^2y}{dt^2} + y = 0:$$

Ենթադրենք ճոճանակի վրա կարելի է ազդել մի u ուժով, որը ուղղված է նրա շարժման գծով և մեծությամբ սահմանափակ է, ասենք՝ $\mid u \mid \leq 1$: Այդ դեպքում u կառավարումով այսպիսի ճոճանակի շարժման հավասարումը կլինի հետևյալը.

$$\frac{d^2y}{dt^2} + y = u$$

(տես գծ 7.1):

Նշանակելով $x_1 = y$, $x_2 = \frac{dy}{dt}$՝ հավասարումը բերենք առաջին կարգի դիֆերենցիալ հավասարումների հետևյալ համակարգի.

$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = -x_1 + u: \end{cases}$$

Այստեղ ֆազային վեկտորի $x = (x_1, x_2)$ առաջին կոորդինատը ցույց է տալիս ճոճանակի շեղումը հավասարակշռության դիրքից, իսկ երկրորդը նրա արագությունն է: Հավասարակշռության վիճակը $(0,0)$ կետն է փազային հարթությունում: Ենթադրենք մի ինչ-որ արտաքին ուժի ազդեցությամբ ճոճանակը տեղափոխվել է $x^0 = (x_1^0, x_2^0)$ վիճակի: Ենթադրենք ճոճանակի գտնվելը այդ վիճակում ցանկալի չէ և անհրաժեշտ է արտաքին u ուժի ազդեցության տակ նրան ամենակարճ ժամանակահատվածում նորից տեղաշարժել հավասրակշռության վիճակ: Սա արագագործության խնդիրն է կառավարվող ֆիզիկական ճոճանակի համակարգում:

7.2 Որոշ տեղեկություններ դիֆերենցիալ հա– վասարումների տեսությունից

Ա) **Սկալյար դեպք**: Դիտարկենք կոշու խնդիրը գծային դիֆերենցիալ հետևյալ հավասարման համար.

$$\frac{dx}{dt} = ax + u, \ x(t_0) = x^0, \tag{7.2.1}$$

որտեղ $x(t)$-ն սկալյար ֆունկցիա է t արգումենտով, a-ն հաստատուն է, իսկ $u(t)$-ն անընդհատ ֆունցիա է: Հայտնի է, որ այս խնդրի լուծումը արտահայտվում է հետևյալ բանաձևով`

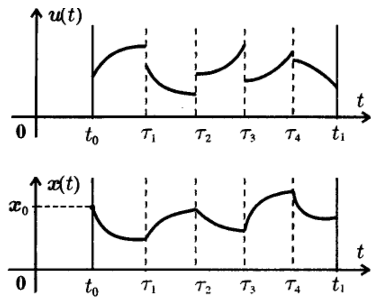
$$x(t) = e^{(t-t_0)a}(x^0 + \int\limits_{t_0}^{t} e^{-(s-t_0)a}u(s)ds): \tag{7.2.2}$$

Դրանում կարելի է համոզվել անմիջապես տեղադրումով: Իրոք պարզ է, որ $x(t_0) = x^0$: Ունենք նաև

$$\frac{dx}{dt} = ae^{(t-t_0)a}(x^0 + \int\limits_{t_0}^{t} e^{-(s-t_0)a}u(s)ds) +$$

$$+e^{-(t-t_0)a}e^{(t-t_0)a}u(t) = ax(t) + u(t):$$

(7.2.2) բանաձևը կոչվում է Կոշիի բանաձև: Եթե $u(t)$-ն կտոր առ կտոր անընդհատ ֆունկցիա է, իսկ $\tau_1, \tau_2, \ldots, \tau_s$-ը՝ նրա խզման կետերը, ապա $x(t)$-ն կտոր առ կտոր դիֆերենցելի է և այն բավարարում է (7.2.1) դիֆերենցիալ հավասարմանը ժամանակի բոլոր $t \neq \tau_i$, $(i = 1, 2, \ldots s)$ պահերին (տես գծ.7.2):



Գծ. 7.2: Հավասարման կտոր առ կտոր դիֆ. լուծումը

Բ) **Ընդհանուր դեպք** ($n > 1$)**:** Դիտարկենք Կոշիի խնդիրը գծային դիֆերենցիալ հավասարումների հետևյալ համակարգի համար.

$$\frac{dx}{dt} = Ax + u, \; x(t_0) = x^0, \tag{7.2.3}$$

որտեղ

$$x = \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix}, A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix}, u(t) = \begin{pmatrix} u_1(t) \\ . \\ . \\ . \\ u_n(t) \end{pmatrix}:$$

(7.2.3) խնդրի լուծումը արտահայտվում է հետևյալ բանաձևով.

$$x(t) = e^{(t-t_0)A}x^0 + \int\limits_{t_0}^{t} e^{(t-s)A}u(s)ds, \qquad (7.2.4)$$

որը կոչվում է Կոշու բանաձև, իսկ $e^{(t-t_0)A}$-ն էքսպոնենցիալ մատրից: Այն սահմանվում է հետևյալ կերպ.

$$e^{tA} = E + \frac{t}{1!}A + \frac{t^2}{2!}A^2 + \dots + \frac{t^k}{k!}A^k + \dots = \sum_0^\infty \frac{t^k}{k!}A^k: \quad (7.2.5)$$

Ցույց տանք, որ այս շարքը ֆիքսած $t \in R^1$ դեպքում բացարձակ զուգամետ է: Իրոք, եթե a_{ij}-ին A մատրիցի էլեմենտ է, ապա $\mid a_{ij} \mid \leq \|A\|$, որտեղ

$$\|A\| = \max_{x \in B_1(0)} \|Ax\|:$$

Մատրիցային (7.2.5) շարքի կամայական p էլեմենտ ներկայացվում է թվային շարքի տեսքով.

$$p = p_1 + p_2 + \dots + p_k + \dots,$$

որտեղ p_k-ն

$$\frac{t^k}{k!}A^k$$

մատրիցի համապատասխան էլեմենտն է: Այսպիսով,

$$\mid p_k \mid \leq \| \frac{t^k}{k^1} A^k \| \leq \frac{\mid t \mid^k}{k!} \|A\|^k,$$

և հետևաբար մատրիցային (7.2.5) շարքը բացարձակ զուգամետ է:

Էքսպոնենցիալ e^{tA} մատրիցը ունի հետևյալ կարևոր հատկությունները.

(ա) $e^{tA}e^{sA} = e^{(t+s)A}$:

Իրոք, քանի որ բացարձակ զուգամետ շարքերը կարելի է բազմապատկել, ապա

$$e^{tA}e^{sA} = (E+tA+\frac{t^2}{2!}A^2+\dots)(E+sA+\frac{s^2}{2!}A^2+\dots) =$$

$$= E + (t+s)A + \frac{1}{2!}(t^2+2ts+s^2)A^2+$$

$$+\frac{1}{3!}(t^3+3t^2s+3ts^2+s^3)A^3+\dots = e^{(t+s)A}\text{:}$$

(բ) $(e^{tA})^T = e^{tA^T}$:

(գ) $\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A$:

Ցույց տանք գ) հատկությունը: Նախ e^{tA} մատրիցի յուրաքանչյուր $(e^{tA})_{ij}$ էլեմենտ անընդհատ դիֆերենցելի ֆունկցիա է, քանի որ

$$(e^{tA})_{ij} = \sum_{k=0}^{\infty} \frac{t^k}{k!}(A^k)_{ij}\text{:}$$

Ռւսփի ասփիճանային շարքը անդամ առ անդամ ածանցելով՝ կունենանք

$$\frac{d}{dt}e^{tA} = \frac{d}{dt}(E + tA + \frac{t^2}{2!}A^k + \dots + \frac{t^k}{k!}A^k + \dots) =$$

$$= A + tA^2 + \dots + \frac{t^{k-1}}{(k-1)!}A^k + \dots =$$

$$= A(E + tA + \dots + \frac{t^{k-1}}{(k-1)!}A^{k-1} + \dots) = Ae^{tA} = e^{tA}A:$$

Այս հափկությունը թույլ է փալիս անմիջականորեն սփուգել, որ (7.2.4) բանաձևով փրվող ֆունկցիան Կոշու խնդրի լուծումն է: Իրոք

$$\frac{dx}{dt} = Ae^{(t-t_0)A}(x^0 + \int\limits_{t_0}^{t} e^{-(s-t_0)A}u(s)ds) +$$

$$+ e^{-(t-t_0)A}e^{(t-t_0)A}u(t) = Ax(t) + u(t):$$

Բերենք էքսպոնենցիալ մափրիցների հաշվման մի քանի օրինակներ:

Օրինակ: Գփնել e^{tA}, եթե $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$:

Պարզ հաշվարկով համոզվում ենք, որ $A^2 = 0$: Հեփևաբար, $A^k = 0,\ k \geq 2$:

Ուրեմն

$$e^{tA} = E + tA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + t\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}:$$

Օրինակ: Գտնել e^{tA}, եթե

$$A=\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}:$$

Հեշտ է տեսնել, որ

$$A^2=-E,\ A^6=-E,$$

$$A^3=-A,\ A^7=-A,$$

$$A^4=E,\ A^8=E,$$

$$A^5=A,\ A^9=A,$$

$$\cdots\ \cdots$$

Հետևաբար

$$e^{tA}=E+tA+\frac{t^2}{2!}(-E)+\frac{t^3}{3!}(-A)+\frac{t^4}{4!}E+$$

$$+\frac{t^5}{5!}A+\frac{t^6}{6!}(-E)+\cdots=$$

$$=(1-\frac{t^2}{2!}+\frac{t^4}{4!}+\dots)E+(t-\frac{t^3}{3!}+\frac{t^5}{5!}+\dots)A=$$

$$=\cos(t)E+\sin(t)A:$$

Այսպիսով,

$$e^{tA}=cost\begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix}+sint\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}=$$

$$\begin{pmatrix} cost & sint\\ -sint & cost\end{pmatrix}, e^{tA^T}=\begin{pmatrix} cost & -sint\\ sint & cost\end{pmatrix}:$$

Օրինակ: Լուծել դիֆերենցիալ հավասարումների հետևյալ համակարգը.

$$\begin{cases} \frac{dx_1}{dt} = x_2 + u_1(t), \\ \frac{dx_2}{dt} = -x_1 + u_2(t) \end{cases}$$

$x(0) = (1, 0)$ սկզբնական պայմանով, իսկ $u(t)$- ն $[0, \pi]$ հատվածի վրա որոշված հետևյալ վեկտոր ֆունկցիան է.

$$u(t) = (u_1(t), u_2(t)) = \begin{cases} (0, 1), \text{եթե } t \in [0, \pi/2] \\ (-1, 0), \text{եթե } t \in [\pi/2, \pi]: \end{cases}$$

Ըստ Կոշիի բանաձևի, եթե $t, \in [0, \pi/2]$, ապա

$$x(t) = \begin{pmatrix} cost & sint \\ -sint & cost \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} +$$

$$+ \int_0^t \begin{pmatrix} cos(t-s) & sin(t-s) \\ -sin(t-s) & cos(t-s) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} ds =$$

$$= \begin{pmatrix} cost \\ -sint \end{pmatrix} + \int_0^t \begin{pmatrix} sin(t-s) \\ cos(t-s) \end{pmatrix} ds =$$

$$= \begin{pmatrix} cost \\ -sint \end{pmatrix} + \begin{pmatrix} 1 - cost \\ sint \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}:$$

Իսկ, եթե $t \in [\pi/2, \pi]$, ապա

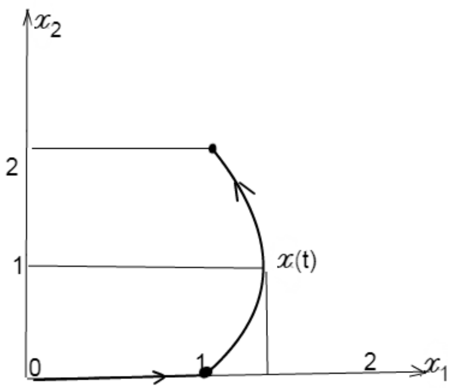
$$x(t) = \begin{pmatrix} sint + cost \\ 1 + cost - sint \end{pmatrix}:$$

Այսպեղից կստանանք

$$x_1^2 + (x_2 - 1)^2 = 2,$$

որը $(0, 1)$ կենտրոնով և $\sqrt{2}$ շառավղով շրջանագիծ է: Այսպիսով, օբեկտը մինչև $\pi/2$ պահը չի շարժվում՝ գտնվելով սկզբնական $x^0 = (1, 0)$ վիճակում: Այնուհետև $[\pi/2, \pi]$ ժամանակահատվածում կատարում է շրջանագծային շարժում՝ շարժվելով վերը նշված շրջանագծով ժամ սլաքի հակառակ ուղղությամբ: Շրջանագծային աղեղի մի կտոր պատկերված է գծագիր 7.3-ում:

Նշենք նաև, որ համակարգի $x(t)$ լուծումը կտոր առ կտոր ողորկ է, քանի որ նրա ածանցյալը խզվում է $t = \pi/2$ պահին:



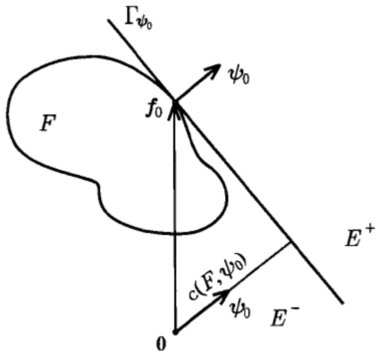
Գծ. 7.3: Համակարգի կտոր առ կտոր դիֆ. հետագիծը

7.3 Հենման ֆունկցիաներ

Ուռուցիկ կոմպակտ բազմությունները նկարագրվում են հենման ֆունկցիաների օգնությամբ:

Սահմանում 7.3.1: *Դիցուք F-ը կոմպակտ ենթաբազմություն է R^n պարածությունից: Այդ բազմության հենման ֆունկցիան սահմանվում է հետևյալ հավասարությամբ.*

$$C(F, \psi) = \max_{f \in F}(f, \psi):$$



Գծ. 7.4: Հենման ֆունկցիայի երկրաչափական մեկնաբանությունը. հենման հիպերհարթությունը պարածությունը բաժանում է երկու կիսապարածությունների: Հենման ֆունկցիայի արժեքը 0 սկզբնակետի հեռավորությունն է այդ հիպերհարթությունից վերցրած + կամ − նշաններով՝ կախված թե այդ կետը դրական կիսապարածությունում է, թե բացասական:

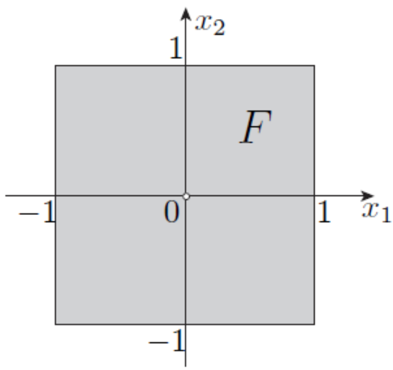
Օրինակ: Դիցուք $F = \{x \in R^2 / x_1^2 + x_2^2 \leq 1\}$:
Այդ դեպքում

$$C(F, \psi) = \|\psi\| = \sqrt{\psi_1^2 + \psi_2^2} :$$

Օրինակ: Դիցուք

$$F = \{x \in R^2 / \mid x_1 \mid \leq 1, \mid x_2 \mid \leq 1\}$$

(տես գծ.7.5):



Գծ. 7.5: Քառակուսի

Այդ դեպքում

$$C(F, \psi) = \max_{|f_1| \leq 1, |f_2| \leq 1} (\psi_1 f_1 + \psi_2 f_2) = \mid \psi_1 \mid + \mid \psi_2 \mid$$

Ուռուցիկ կոմպակտ բազմությունը ներկայացվում է նրա հենման ֆունկցիայի միջոցով: Տեղի ունի հետևյալ պնդումը:

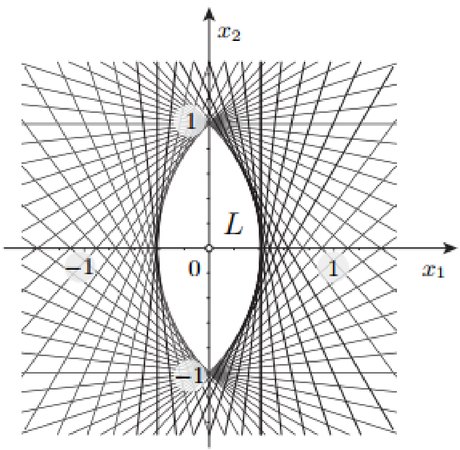
Թեորեմ 7.3.1: *Դիցուք* F*-ը ուռուցիկ կոմպակտ ենթաբազմություն է* R^n *տարածությունից: Այդ դեպքում տեղի ունի հետևյալ ներկայացումը.*

$$F = \bigcap_{\psi \in S_1(0)} \{f \in R^n / (f, \psi) \leq C(F, \psi)\}: \qquad (7.3.1)$$

Այստեղ

$$G_\psi \equiv \{f \in R^n / (f, \psi) \leq C(F, \psi)\}$$

բազմությունը կիսատարածություն է, որը ծնվում է ψ *նորմալով* F *բազմության համար կառուցված հենման հիպերհարթությամբ (տես գծ. 7.6):*



Գծ. 7.6: Ուռուցիկ կոմպակտի ներկայացումը

► Դիցուք $f \in F$: Այդ դեպքում, ըստ հենման ֆունկցիայի սահմանման, կամայական $\psi \in S_1(0)$ վեկտորի համար կունենանք՝

$$(f,\psi) \le \max_{g\in F}(g,\psi) = C(F,\psi) :$$

Հետևաբար f-ը պատկանում է (7.3.1) հավասարության աջ մասին: Այժմ ենթադրենք, որ $f_0 \notin F$: Այդ դեպքում, ըստ ուռուցիկ բազմությունների խիստ բաժանման թեորեմի, գոյություն կունենա այնպիսի $\psi_0 \in S_1(0)$, որ

$$C(F,\psi_0) < (f_0,\psi_0),$$

ինչը նշանակում է, որ f_0-ն չի պատկանում (7.3.1) հավասարության աջ մասին: ■

Հետևանք 7.3.1: *Դիցուք $F, G \subset R^n$ բազմությունները ուռուցիկ են և կոմպակտ: Այդ դեպքում*

$$F = G \Longleftrightarrow C(F,\psi) = C(G,\psi)\ \forall\psi \in S_1(0):$$

► $C(F,\psi) = C(G,\psi)\ \forall\psi \in S_1(0) \Rightarrow$

$$\Rightarrow (f,\psi) \le C(G,\psi)\ \forall f \in F,\ \psi \in S_1(0) \Rightarrow$$

$$\Rightarrow f \in G \Rightarrow F \subseteq G :$$

Անալոգ ձևով ստանում ենք

$$G \subseteq F:$$

Հետևաբար $F = G$: ■

Օրինակ: Գտնել $F_1 = B_{r_1}(a_1),\ \ F_2 = B_{r_2}(a_2)$ գնդերի հանրահաշվական գումարը՝ $F = F_1 + F_2$: Հաշվենք F բազմության հենման ֆունկցիան.

$$C(F,\psi) = C(F_1 + F_2,\psi) = C(F_1,\psi) + C(F_2,\psi) =$$

$$= (a_1 + a_2, \psi) + (r_1 + r_2)\|\psi\| = C(B_{r_1+r_2}(a_1 + a_2), \psi):$$

Հետևաբար ըստ նախորդ թեորեմի՝

$$F = B_{r_1+r_2}(a_1 + a_2):$$

Հետագայում կառավարման պայմանի ստացման ժամանակ կարևոր դեր է խաղալու հետևյալ պնդումը:

Թեորեմ 7.3.2: *Երկու ուռուցիկ կոմպակտ* F, G *բազմություններ հատվում են այն և միայն այն դեպքում, երբ*

$$C(F, \psi) + C(G, -\psi) \geq 0 \; \forall \psi \in S_1(0):$$

▶ $F \cap G \neq \emptyset \Leftrightarrow$

$$\Leftrightarrow 0 \in F + (-D) \Leftrightarrow C(F + (-G), \psi) \geq 0 \; \forall \psi \Leftrightarrow$$

$$\Leftrightarrow C(F, \psi) + C(G, -\psi) \geq 0 \; \forall \psi \in S_1(0):$$

Նշենք, որ ապացույցի ընթացքում կիրառվեց այն պնդումը, որ երկու ուռուցիկ կոմպակտ բազմությունների գումարի հենման ֆունկցիան հավասար է բազմությունների հենման ֆունկցիաների գումարին: Այս փաստը խնդրի տեսքով առաջարկվում է ներքևում (տես խնդիր 6):

ԽՆԴԻՐՆԵՐ

1. Դիցուք $G \subset R^n$ կոմպակտ բազմություն է: Ապացուցել, որ
$$C(G, \psi) = C(conv(G), \psi) \; \forall \psi \in R^n :$$

2. Դիցուք $G \subset R^n$ կոմպակտ բազմություն է, իսկ $A(n \times n)$-ը մատրից է: Ապացուցել, որ
$$C(AG, \psi) = C(G, A^T \psi) \; \forall \psi \in R^n:$$

3. Գտնել ուղղանկյան հենման ֆունկցիան, որի կողմերը զուգահեռ են կոորդինատային առանցքներին.

$$M = \{x \in R^2 / a_1 \leq x_1 \leq b_1,\ a_2 \leq x_2 \leq b_2\}:$$

4. Գտնել F բազմության հենման ֆունկցիան, որը ներկայացվում է հետևյալ բանաձևով.

$$F = \{x \in R^2 / \mid x_1 + x_2 \mid \leq 1,\ \mid x_1 - x_2 \mid \leq 1\}:$$

5. Գտնել էլիպսի հենման ֆունկցիան, որը հարթության վրա տրվում է հետևյալ բանաձևով.

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 :$$

6. Դիցուք $F, G \subset R^n$ ուռուցիկ կոմպակտ բազմություններ են: Ապացուցել, որ

$$C(F+G, \psi) = C(F, \psi) + C(G, \psi)\ \forall \psi \in R^n:$$

7.4 Հասանելիության բազմություն

Դիտարկենք գծային հավասարումների հետևյալ համակարգը.

$$\frac{dx}{dt} = Ax + u:$$

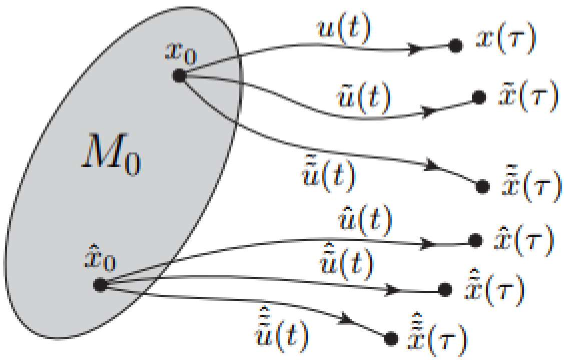
Որպես թույլատրելի կառավարումների բազմություն վերցնենք ըստ Լեբեգի ինտեգրելի ֆունկցիաների դասը:

Հասանելիության բազմություն կոչվում է փազային տարածության այն բոլոր կետերի բազմությունը, որտեղ հասնում են դիֆերենցիալ հավասարման լուծումների հետագծերը t պահին բոլոր հնարավոր թույլատրելի $u(t) \in U$ կառավարումների օգնությամբ: Այն նշանակենք $X(t)$-ով: Ուստի՝

$$X(t) = \{x(t) = e^{(t-t_0)A}x^0 + \int\limits_{t_0}^{t} e^{(t-s)A}u(s)ds,\ x^0 \in M_0,$$

$$u(s) \in U\}$$

(տես գծ.7.7):



Գծ. 7.7: Հասանելիության բազմությունը

Տեղի ունի հետևյալ կարևոր պնդումը, որը բերում ենք առանց ապացույցի:

Թեորեմ 7.4.1 *[15]: Եթե M_0 և U բազմությունները ուռուցիկ են ու կոմպակտ, ապա ժամանակի կամայական t պահին հասանելիության $X(t)$ բազմությունը նույնպես ուռուցիկ է և կոմպակտ:*

Թեորեմ 7.4.2: *Հասանելիության $X(t)$ բազմության հենման ֆունկցիան ներկայացվում է հետևյալ բանաձևով.*

$$C(X(t),\psi) = C(M_0, e^{(t-t_0)A^T}\psi) + \int\limits_{t_0}^{t} C(U, e^{(t-s)A^T}\psi)ds:$$

► Այս հիմնական թեորեմի ապացույցը հենվում է հետևյալ պնդումների վրա:

Լեմմ 7.4.1: *Դիցուք $U \subset R^n$ կոմպակտ ենթաբազմություն է, իսկ $\eta : [a,b] \to R^n$ անընդհատ վեկտոր ֆունկցիա է:* Այդ դեպքում

$$f(t) \equiv C(U, \eta(t))$$

ֆունկցիան անընդհատ է $[a,b]$-ում:

► Դիցուք $t_1,\ t_2 \in [a,b]$ ֆիքսած կետեր են: Քանի որ U-ն կոմպակտ է, ապա, ըստ Վայերշտրասի թեորեմի, գոյություն ունեն այնպիսի $u_1,\ u_2 \in U$ կետեր, որ

$$(\eta(t_1), u_1) = \max_{u \in U}(u, \eta(t_1)),$$

$$(\eta(t_2), u_2) = \max_{u \in U}(u, \eta(t_2)):$$

Քանի որ U բազմությունը սահմանափակ է, ապա գոյություն կունենա այնպիսի $m > 0$ թիվ, որ

$$\|u\| \leq m \, \forall u \in U :$$

Հետևաբար

$$f(t_2) - f(t_1) = C(U, \eta(t_2)) - C(U, \eta(t_1)) =$$

$$= (\eta(t_2), u_2) - (\eta(t_1), u_1) \leq (\eta(t_2), u_2) - (\eta(t_1), u_2) \leq$$

$$\leq \|\eta(t_2) - \eta(t_1)\| \|u\| \leq m\|\eta(t_2 - \eta(t_1)\|:$$

Անալոգ ձևով կստանանք

$$f(t_1) - f(t_2) = C(U, \eta(t_1)) - C(U, \eta(t_2)) \leq$$

$$\leq m\|\eta(t_2) - \eta(t_1)\|:$$

Այս երկու անհավասարություններից հետևում է, որ

$$| C(U, \eta(t_1)) - C(U, \eta(t_2)) | \leq m\|\eta(t_2) - \eta(t_1)\|,$$

ինչը նշանակում է, որ $f(t)$ ֆունկցիան անընդհատ է $[a, b]$ հատվածի վրա: ■

Լեմմ 7.4.2 *(Չափելի ճյուղի անջատման մասին): Դիցուք $U \subset R^n$ կոմպակտ ենթաբազմություն է, իսկ $\eta : [a, b] \to R^n$ անընդհատ վեկտոր ֆունկցիա է:*

Այդ դեպքում գոյություն կունենա այնպիսի չափելի $u(t) \in U$ $\forall t \in [a, b]$ վեկտոր ֆունկցիա, որ

$$C(U, \eta(t)) = (u(t), \eta(t)) \;\; \forall t \in [a, b]:$$

► Յուրաքանչյուր t-ի համար վերցնենք այնպիսի $v(t)$ վեկտոր

$$\Omega(t) = \{v \in U / C(U, \eta(t)) = (v(t), \eta(t))\}$$

բազմությունից, որի առաջին կոորդինատը ամենափոքրն է: Այդպիսի վեկտոր կա, քանի որ U կոմպակտ է: Եթե նման վեկտորները մի քանի հատ են, ապա դրանց միջից ընտրում ենք այն վեկտորը, որի երկրորդ կոորդինատն ամենափոքրն է, և այսպես շարունակ կառուցում ենք միակ $u(t) \in U$ վեկտորը՝ բավարարող

$$C(U, \eta(t)) = (u(t), \eta(t))$$

պայմանին: Ցույց տանք, որ $u(t)$-ն չափելի է $[a, b]$-ի վրա: Ապացույցը կատարենք ինդուկցիայով: Ենթադրենք $u(t)$ վեկտորի $u_1(t), \ldots, u_{s-1}(t)$ կոմպոնենտները չափելի են: Ապացուցենք, որ $u_s(t)$- ն նույնպես չափելի է: Ըստ Լուզինի թեորեմի ընտրենք փակ $E_l \subseteq [a, b]$ բազմությունների այնպիսի ընտանիք, որ

$$\mu([a, b] \backslash E_l) < 2^{-l}$$

և $u_1(t), \ldots, u_{s-1}(t)$ ֆունկցիաները լինեն անընդհատ E_l $l = 1, 2, \ldots$ բազմությունների վրա: Վերցնենք կամայական α թիվ և ցույց տանք, որ

$$A_\alpha = \{t \in E_l / u_s(t) \leq \alpha\}$$

բազմությունը փակ է, որից և անմիջականորեն կբխի $u_s(t)$-ի չափելիությունը ամբողջ $[a, b]$ հատվածի վրա: Ենթադրենք հակառակը: Այդ դեպքում գոյություն կունենա այնպիսի $\{t_k\} \in E_l, \ t_k \to \bar{t}$ հաջորդականություն, որ

$$u_s(t_k) \leq \alpha < u_s(\bar{t}):$$

Քանի որ U-ն կոմպակտ է, ապա կարող ենք ենթադրել, որ

$$u(t_k) \to \overline{u} \in U:$$

Քանի որ $u_1(t), \ldots, u_{s-1}(t)$ և $\eta(t)$ ֆունկցիաները անընդհատ են E_l բազմության վրա, ապա

$$u_i(t_k) \to \overline{u_i}, \ i = 1, \ldots, s-1,$$

$$\eta(t_k) \to \eta(\bar{t}):$$

Այնպես որ
$$C(U, \eta(\bar{t})) = (\eta(\bar{t}), \overline{u}):$$

Մյուս կողմից
$$\overline{u_s} \leq \alpha < u_s(\bar{t}),$$

որը հակասում է $u_s(\bar{t})$-ի կառուցմանը: ■

Լեմմ 7.4.3: Դիցուք

$$Y_U \equiv \{u(t) \in L_1[a, b]/u(t) \in U, \ \forall t \in [a, b]\}:$$

Այդ դեպքում կամայական ψ-ի համար

$$C(\int_{t_0}^{t} e^{(t-s)A} Y_U ds, \psi) = \int_{t_0}^{t} C(e^{(t-s)A} U, \psi) ds :$$

► Իրոք դիցուք $u(t) \in Y_U$: Այդ դեպքում

$$(\int_{t_0}^{t} e^{(t-s)A} u(s) ds, \psi) = (\int_{t_0}^{t} (e^{(t-s)A} u(s), \psi) ds \leq$$

$$\leq \int_{t_0}^{t} C(e^{(t-s)A} U, \psi) ds :$$

Ուրեմն

$$C(\int_{t_0}^{t} e^{(t-s)A} Y_U ds, \psi) \leq \int_{t_0}^{t} C(e^{(t-s)A} U, \psi) ds \ \forall \psi :$$

Հակառակ անհավասարությունը բխում է նախորդ լեմմայից: Իրոք, ըստ Լեմմ 7.4.2-ի՝ գոյություն ունի այնպիսի $u(t) \in Y_U$, որ

$$\int_{t_0}^{t} C(e^{(t-s)A}U, \psi)ds = \int_{t_0}^{t} (e^{(t-s)A}u(s), \psi)ds =$$

$$= (\int_{t_0}^{t} e^{(t-s)A}u(s)ds, \psi) \leq C(\int_{t_0}^{t} e^{(t-s)A}Y_U ds, \psi):$$

Թեորեմ 7.4.2 ապացույցը: Ըստ Լեմմ 7.4.3՝ ունենք

$$C(X(t), \psi) = C(e^{(t-t_0)}M_0, \psi) + C\int_{t_0}^{t} (e^{(t-s)A}Y_U, \psi)ds =$$

$$= C(e^{(t-t_0)}M_0, \psi) + \int_{t_0}^{t} C(e^{(t-s)A}U, \psi)ds =$$

$$= C(M_0, e^{(t-t_0)A^T}\psi) + \int_{t_0}^{t} C(U, e^{(t-s)A^T}\psi)ds: \blacksquare$$

Օրինակ: Գտնել

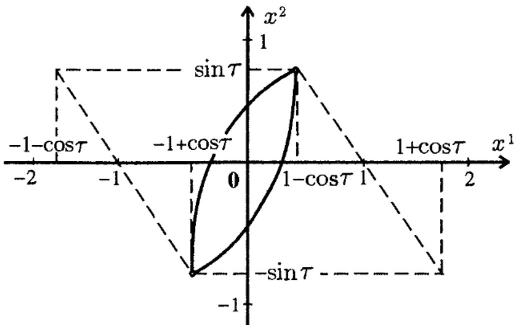
$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = -x_1 + u, \end{cases}$$

համակարգի հասանելիության $X(\pi)$ բազմությունը $t = \pi$ պահին, եթե $M_0 = (0, 0)$ և $\mid u \mid \leq 1$: Նշանակենք

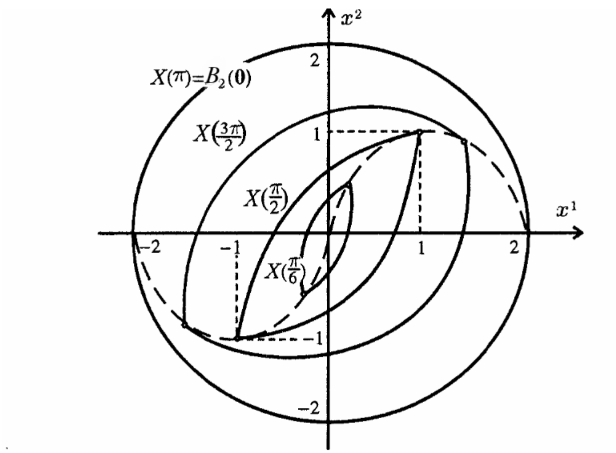
$$U = \{u = (u_1, u_2) \in R^2 / u_1 = 0, \mid u_2 \mid \leq 1\}:$$

Այդ դեպքում
$$C(U,\psi) = \mid \psi_2 \mid:$$
Մենք արդեն գիտենք, որ
$$e^{tA^T} = \begin{pmatrix} cost & -sint \\ sint & cost \end{pmatrix}:$$
Հետևաբար ըստ հասանելիության բազմության հենման ֆունկցիայի բանաձևի՝ կունենանք՝
$$C(X(\pi),\psi) = \int\limits_0^{\pi} \mid \psi_1 sins + \psi_2 coss \mid ds:$$
Այստեղ տեղադրելով $\psi_1 = cos\alpha,\ \psi_2 = sin\alpha$՝ կստանանք
$$C(X(\pi),\psi) = \int\limits_0^{\pi} \mid sin(\alpha + s) \mid ds = \int\limits_{\alpha}^{\alpha+\pi} \mid sin(\theta) \mid d\theta = 2:$$
Այստեղից ստանում ենք $X(\pi) = B_2(0)$:

Գծ.7.8-ում պատկերված է համակարգի հասանելիության բազմությունը τ պահին, որպես
$$B_2(1 + cos\tau, -sin\tau),\ B_2(-1 - cos\tau, sin\tau)$$
շրջանների հատում փազային հարթության վրա: Իսկ գծ.7.9-ում պատկերված է հասանելիության $X(t)$ բազմությունների շարժման դինամիկան $[0,\pi]$ ժամանա–կահատվածում:



Գծ. 7.8: Հասանելիության բազմությունը τ պահին



Գծ. 7.9: Հասանելիության բազմությունների շարժման դինամիկան

ԽՆԴԻՐՆԵՐ

1. Գտնել u կառավարումով
$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = u \end{cases}$$
համակարգի հասանելիության $X(t)$ բազմությունները, եթե $0 \leq u \leq 1,\ M_0 = \{(0,0)\}$:

2. Գտնել u կառավարումով
$$\begin{cases} \frac{dx_1}{dt} = x_2 + u_1, \\ \frac{dx_2}{dt} = -x_1 + u_2 \end{cases}$$
համակարգի հասանելիության $X(t)$ բազմությունները, եթե
$$u = (u_1, u_2) \in B_1((0,0)),\ M_0 = \{x \in R^2 / \mid x_1 \mid \leq 2,\ x_2 = 0\}:$$

7.5 Կառավարման անհրաժեշտ և բավարար պայմանը

Դիտարկենք կառավարվող համակարգ, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով.
$$\frac{dx}{dt} = Ax + u:$$
Թույլատրելի կառավարումը կամայական չափելի $u(t) \in U$ ֆունկցիա է, որտեղ U-ն ուռուցիկ կոմպակտ ենթաբազմություն է R^n տարածությունից, իսկ դիֆերենցիալ

հավասարման լուծումը կամայական բացարձակ անընդհատ ֆունկցիա է, որը համարյա ամենուրեք բավարարում է դիֆերենցիալ հավասարմանը: Դիցուք տրված են սկզբնական և վերջնական ուռուցիկ և կոմպակտ M_0, M_1 բազմությունները: Կառավարման խնդիրն է. **գոյություն ունի արդյո՞ք թույլատրելի $u(t)$ կառավարում, որոշված ինչ որ մի վերջավոր $[t_0, t_1]$ հատվածի վրա, որ նրան համապատասխան $x(t)$ հետագիծը բավարարի**

$$x(t_0) \in M_0, \ x(t_1) \in M_1$$

պայմաններին:

Սահմանենք $\varphi(\psi)$ ֆունկցիան հետևյալ բանաձևով.

$$\varphi(\psi) = C(M_0, e^{(t_1-t_0)A^T}\psi) + C(M_1, -\psi) +$$

$$+ \int\limits_0^{t_1-t_0} C(U, e^{sA^T}\psi)ds: \tag{7.5.1}$$

Թեորեմ 7.5.1 *(կառավարման մասին): Դիցուք տեղի ունեն վերը նշված պայմանները: Որպեսզի համակարգը լինի կառավարելի $[t_0, t_1]$ հատվածի վրա անհրաժեշտ է և բավարար, որ*

$$\varphi(\psi) \geq 0,$$

կամ որ նույնն է՝

$$\min_{\psi \in B_1(0)} \varphi(\psi) \geq 0:$$

► Ակնհայտ է, որ օբյեկտը կլինի կառավարելի $[t_0, t_1]$ միջակայքում միայն այն դեպքում, երբ

$$X(t_1) \bigcap M_1 \neq \emptyset \ :$$

Այս պայմանը համարժեք է հետևյալին՝

$$C(X(t_1),\psi)+C(M_1,-\psi)\geq 0\ \forall\psi\in B_1(0):$$

Տեղադրելով այստեղ հասանելիության $X(t_1)$ բազմության հենման ֆունկցիայի բանաձևը՝ կստանանք վերը նշված անհրաժեշտ ու բավարար պայմանը: ■

Օրինակ: Դիտարկենք ֆիզիկական ճոճանակի շարժման հավասարումը.

$$\begin{cases}\frac{dx_1}{dt}=x_2,\\ \frac{dx_2}{dt}=-x_1+u,\end{cases},$$

$$\mid u\mid\leq 1,\ M_0=\{x\in R^2/x_1=-5,\ \mid x_2\mid\leq 1\},$$

$$M_1=\{(0,0)\}:$$

$M_0,\ M_1,\ U$ բազմությունների հենման ֆունկցիաները ունեն հետևյալ տեսքերը.

$$C(M_0,\psi)=-5\psi_1+\mid\psi_2\mid,\ C(M_1,\psi)=0,$$

$$C(U,\psi)=\mid\psi_2\mid:$$

Ունենք նաև

$$e^{tA^T}=\begin{pmatrix}cost & -sint\\ sint & cost\end{pmatrix},\ \psi=(cos\alpha,sin\alpha):$$

Հետևաբար

$$\varphi(\psi)=\varphi(\alpha)=$$

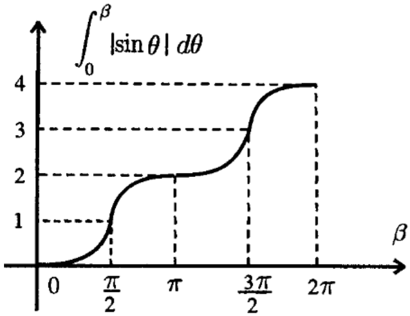
$$=-5cos(\alpha-\tau)+\mid sin(\alpha-\tau)\mid+$$

$$+\int\limits_0^\tau\mid sin(\alpha-s)\mid ds\geq 0,$$

եթե $\tau = t_1 - t_0 \geq 3\pi$: Իրոք, կարելի է ցույց տալ, որ

$$\int_0^{\tau} \mid sin(\theta) \mid d\theta = 2[\frac{\tau}{\pi}] + 1 - cos(\tau - [\frac{\tau}{\pi}]\pi)$$

(տես գծ.7.10): Հետևաբար համակարգը կառավարելի է կամայական $[t_0, t_1]$ հատվածում, որի $\tau = t_1 - t_0$ երկարությունը մեծ է 3π:



Գծ. 7.10:

Օրինակ: Դիտարկենք հետևյալ համակարգը.

$$\begin{cases} \frac{dx_1}{dt} = x_2 + u_1, \\ \frac{dx_2}{dt} = -x_1 + u_2, \end{cases},$$

$$u = (u_1, u_2) \in B_1((0,0)),\ M_0 = B_\pi((3\pi, 0)),\ M_1 = B_\pi(0):$$

Անհրաժեշտ է գտնել $[t_0, t_1]$ հատված, որում համակարգը կառավարելի է: Դիցուք $\tau = t_1 - t_0$:

$U,\ M_0, M_1$ բազմությունների հենման ֆունկցիաներն են՝

$$C(U, \psi) = \|\psi\|,\ C(M_0, \psi) = 3\pi\psi_1 + \pi\|\psi\|,$$

$$C(M_1, \psi) = \pi\|\psi\|:$$

Տեղադրելով այս բանաձևերը կառավարման φ ֆունկցիայի բանաձևի մեջ՝ կստանանք

$$\varphi(\psi) = 3\pi(\psi_1 cos\tau - \psi_2 sin\tau) + 2\pi\|\psi\| + \tau\|\psi\|:$$

Այստեղից,

$$\min_{\psi \in B_1(0)} \varphi(\psi) = \tau - \pi:$$

Հետևաբար համակարգը կառավարելի է ամեն մի հատվածի վրա, որի երկարությունը մեծ կամ հավասար է π, և կառավարելի չէ յուրաքանչյուր π-ից փոքր երկարություն ունեցող հատվածի վրա:

Այժմ որոշենք համակարգի հասանելիության $X(t)$ բազմության շարժման դինամիկան ժամանակի ընթացքում: Տեղադրելով բազմությունների հենման ֆունկցիաների բանաձևերը հասանելիության բազմության հենման ֆունկցիայի բանաձևի մեջ՝ կստանանք

$$C(X(t), \psi) = 3\pi cos\tau\psi_1 - 3\pi sin\tau\psi_2 + (\pi + \tau)\|\psi\| =$$

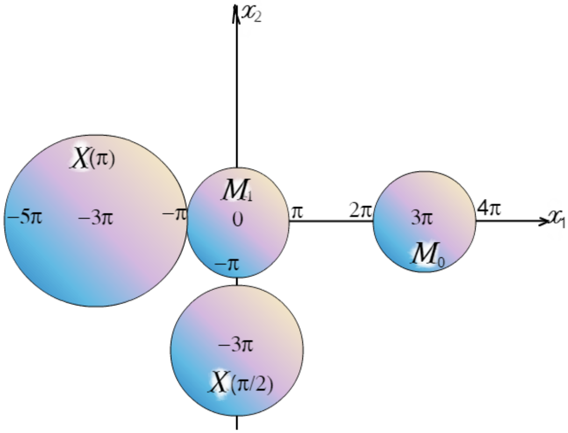
$$= C(B_{r(t)}(a(t)), \psi),$$

որտեղ $r(t) = \pi + t,\ a(t) = 3\pi \begin{pmatrix} cost \\ -sint \end{pmatrix}:$

Հետևաբար հասանելիության $X(t)$ բազմությունը $r(t)$ շառավղով և $a(t)$ կենտրոնով շրջան է, որը ժամանակի ընթացքում պտտվում է 0 կենտրոնով և 3π շառավղով շրջանագծով ժամ սլաքի ուղղությամբ: Ակնհայտ է, որ

$$X(0) = B_\pi((3\pi, 0)) = M_0,\ X(\pi/2) = B_{3\pi/2}((0, -\pi)),$$

$$X(\pi) = B_\pi((-3\pi, 0)):$$



Գծ. 7.11: Հասանելիության բազմության դինամիկան $t = 0, \pi/2, \pi$ պահերին

Գծ.7.11-ում պատկերված է հասանելիության $X(t)$ բազմության շարժման դինամիկան $[0, \pi]$ ժամանակահատվածում:

ԽՆԴԻՐՆԵՐ

---

1. Դիտարկենք օբյեկտ, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով.

$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = -x_1 + u, \end{cases}$$

որտեղ

$$u \in [0, 1], M_0 = \{(2, 0)\},\ M_1 = \{(0, 0)\}:$$

Ցույց տալ, որ օբյեկտը կառավարելի է $[0, 2\pi]$ հատվածի վրա:

2. Արդյո՞ք կառավարելի է օբյեկտը $[0, 3]$ ժամանակահատվածում, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով.

$$\begin{cases} \frac{dx_1}{dt} = u_1, \\ \frac{dx_2}{dt} = -x_1 + u_2, \end{cases},$$

որտեղ $\mid u_1 \mid \leq 1, \ \mid u_2 \mid \leq 1, \ M_0 = \{(-3, 4)\}, \ M_1 = \{(0, 0)\}$:

7.6 Պոնտրյագինի մաքսիմումի սկզբունքը օպտիմալ արագագործության գծային խնդրում

Դիտարկենք արագագործության խնդիրը գծային համակարգերում.

$$\frac{dx}{dt} = Ax + u, \ u(t) \in U,$$

$$x(t_0) \in M_0, \ x(t_1) \in M_1, \ T = t_1 - t_0 \to min:$$

Լեմմ 7.6.1: *Դիցուք $T = t_1 - t_0$-ը օպտիմալ ժամանակամիջոցն է: Այդ դեպքում գոյություն ունի համալուծ*

$$\frac{d\psi}{dt} = -A^T \psi$$

համակարգի այնպիսի ոչ տրիվիալ $\psi(t)$ լուծում, որ

$$C(X(t_1), \psi(t_1)) + C(M_1, -\psi(t_1)) = 0:$$

▶ Քանի որ T-ն օպտիմալ ժամանակամիջոցն է, ապա

$$X(t_1)\bigcap M_1 \neq \emptyset, \tag{7.6.1}$$

$$X(t)\bigcap M_1 = \emptyset \; \forall t \in [t_0, t_1): \tag{7.6.2}$$

(7.6.1) պայմանը համարժեք է հետևյալին.

$$C(X(t_1), \psi) + C(M_1, -\psi) \geq 0, \; \forall \psi \in S_1(0): \tag{7.6.3}$$

Դիտարկենք $t_0 < t_k < t_1, \; t_k \to t_1$ հաջորդականությունը: (7.6.2) պայմանից հետևում է, որ

$$X(t_k)\bigcap M_1 = \emptyset:$$

Այսինքն՝ գոյություն կունենա այնպիսի $p_k \in S_1(0)$ հաջոր–դականություն, որ

$$C(X(t_k), p_k) + C(M_1, -p_k) < 0, \; \forall k: \tag{7.6.4}$$

Քանի որ $p_k \in S_1(0)$, ապա ընդհանրությունը չխախտելով, կարող ենք ենթադրել, որ $p_k \to p_*$:
(7.6.4) անհավասարության մեջ անցնելով սահմանի, երբ $k \to \infty$ կստանանք՝

$$C(X(t_1), p_*) + C(M_1, -p_*) \leq 0 : \tag{7.6.5}$$

(ապացույցը տես [9]-ում, Թեորեմ 1, էջ 59):

Նկատի ունենալով նաև (7.6.3) անհավասարությունը՝ կունենանք

$$C(X(t_1), p_*) + C(M_1, -p_*) = 0:$$

Եթե վերցնենք համալուծ համակարգի լուծումը $\psi(t_1) = p_*$ պայմանով՝ կստանանք լեմմի պնդումը: ■

Թեորեմ 7.6.1 *(Պոնտրյագինի մաքսիմումի սկզբունքը)* : *Դիցուք* $(u(t), x(t))$*-ն օպտիմալ պրոցես է արագագործության խնդրում: Այդ դեպքում գոյություն ունի համալուծ համակարգի այնպիսի ոչ տրիվիալ* $\psi(t)$ *լուծում, որ տեղի ունի*

1. *մաքսիմումի պայմանը.*

$$(u(t), \psi(t)) = C(U, \psi(t)) \; \forall t \in [t_0, t_1]:$$

2. *Տրանսվերսալության պայմանը* M_0 *բազմության վրա.*

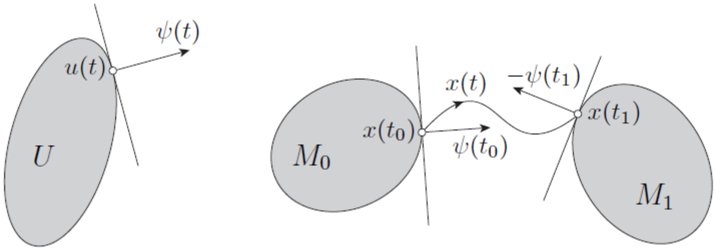
$$(x(t_0), \psi(t_0)) = C(M_0, \psi(t_0)):$$

Այս պայմանը նշանակում է, որ $\psi(t_0)$ *վեկտորը հենման վեկտոր է* M_0 *բազմության համար* $x(t_0)$ *կետում:*

3. *Տրանսվերսալության պայմանը* M_1 *բազմության վրա.*

$$(x(t_1), -\psi(t_1)) = C(M_1, -\psi(t_0)):$$

Այսինքն՝ $-\psi(t_1)$ *վեկտորը* M_1 *բազմության հենման վեկտորն է* $x(t_1)$ *կետում (տես 7.12):*



Գծ. 7.12: Տրանսվերսալության պայմանները

► Ըստ Լեմմ 7.6.1-ի՝ գոյություն ունի համալուծ համակարգի այնպիսի ոչ տրիվիալ $\psi(t)$ լուծում, որ

$$C(X(t_1), \psi(t_1)) + C(M_1, -\psi(t_1)) = 0 :$$

Այս հավասարությունից անդամ առ անդամ հանելով ակնհայտ

$$(x(t_1), \psi(t_1)) + (x(t_1), -\psi(t_1)) = 0$$

հավասարությունը՝ կստանանք

$$[C(X(t_1), \psi(t_1)) - (x(t_1), \psi(t_1))] +$$

$$[C(M_1, -\psi(t_1)) - (x(t_1), -\psi(t_1))] = 0:$$

Այս գումարի երկու գումարելիները ոչ բացասական են, քանի որ

$$x(t_1) \in X(t_1),\ x(t_1) \in M_1:$$

Հետևաբար

$$C(X(t_1), \psi(t_1)) - (x(t_1), \psi(t_1)) = 0, \tag{7.6.6}$$

$$C(M_1, -\psi(t_1)) - (x(t_1), -\psi(t_1)) = 0: \qquad (7.6.7)$$

(7.6.7)-ը փրանսվերսալության պայմանն է M_1 բազմության վրա: Ցույց տանք, որ (7.6.6) պայմանից հետևում է մաքսիմումի սկզբունքը և փրանսվերսալության պայմանը M_0 բազմության վրա: Համաձայն Կոշիի բանաձևի՝ ունենք

$$x(t) = e^{(t-t_0)A}x(t_0) + \int_{t_0}^{t_1} e^{(t-s)A}u(s)ds : \qquad (7.6.8)$$

Քանի որ համալուծ համակարգի լուծումը ունի

$$\psi(t) = e^{(t_0-t)A^T}\psi(t_0)$$

տեսքը, ապա (7.6.8) պայմանից կստանանք՝

$$(x(t), \psi(t)) = (e^{(t-t_0)A}x(t_0), e^{(t_0-t)A^T}\psi(t_0)) +$$

$$+ \int_{t_0}^{t} (e^{(t-s)A}u(s), e^{(t_0-t)A^T}\psi(t_0))ds =$$

$$= (x(t_0), \psi(t_0)) + \int_{t_0}^{t} (u(s), e^{(t_0-s))A^T}\psi(s)ds =$$

$$= (x(t_0), \psi(t_0)) + \int_{t_0}^{t} (u(s), \psi(s))ds: \qquad (7.6.9)$$

Տեղադրելով հասանելիության բազմության հենման ֆունկցիայի բանաձևի մեջ՝ ψ-ի փոխարեն համալուծ համակարգի $\psi(t)$ լուծումը՝ կստանանք

$$C(X(t), \psi(t)) = C(M_0, e^{(t-t_0)A^T}e^{(t_0-t)A^T}\psi(t_0)) +$$

$$+\int_{t_0}^{t} C(U, e^{(t-s)A^T} e^{(t_0-t)A^T}\psi(t_0))ds =$$

$$= C(M_0, \psi(t_0)) + \int_{t_0}^{t} C(U, \psi(s))ds: \qquad (7.610)$$

(7.6.6) հավասարության մեջ տեղադրելով (7.6.9) և (7.6.10) բանաձևերը՝ կստանանք

$$[C(X(t_0), \psi(t_0)) - (x(t_0), \psi(t_0))]+$$

$$+\int_{t_0}^{t} [C(U, \psi(s)) - (u(s), \psi(s))]ds = 0:$$

Այս հավասարության երկու գումարելիներն էլ ոչ բացասական են հենման ֆունկցիայի սահմանման համաձայն: Ուրեմն

$$C(X(t_0), \psi(t_0)) - (x(t_0), \psi(t_0)) = 0, \qquad (7.6.11)$$

$$\int_{t_0}^{t} [C(U, \psi(s)) - (u(s), \psi(s))]ds = 0: \qquad (7.6.12)$$

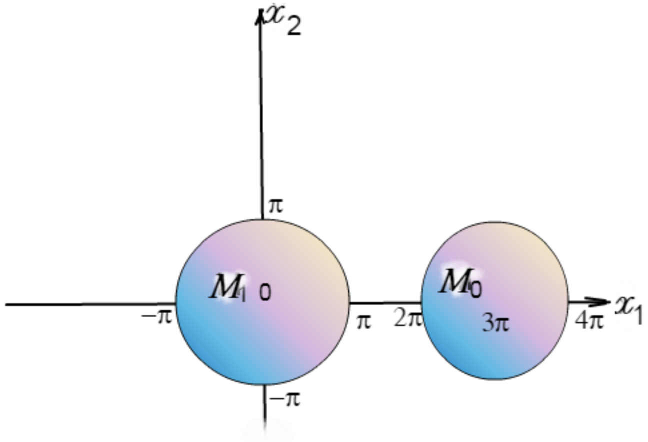
(7.6.11) պայմանը տրանսվերսալության պայմանն է M_0 բազմության վրա, իսկ (7.6.12) պայմանից անմիջապես բխում է մաքսիմումի սկզբունքը.

$$C(U, \psi(t)) = (u(t), \psi(t)) \ \forall t \in [t_0, t_1]: \blacksquare$$

Օրինակ: Գտնել օպտիմալ կառավարումը և օպտիմալ հետագիծը արագագործության հետևյալ խնդրում.

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ U = B_1(0),$$

$$M_0 = B_\pi((3\pi, 0)),\ M_1 = B_\pi(0) \text{ (տես. գծ. 7.13):}$$



Գծ. 7.13: Սկզբնական և նպատակային բազմությունները

Օպտիմալ$(u(t), x(t)),\ t \in [0, t_1],\ x(t_0) \in M_0,\ x(t_1) \in M_1$ զույգը բավարարում է մաքսիմումի սկզբունքին.

1. $(u(t), \psi(t)) = C(U, \psi(t)),\ t \in [0, t_1],$
2. $(x(t_0), \psi(t_0)) = C(M_0, \psi(t_0)),$
3. $(x(t_1), -\psi(t_1)) = C(M_1, -\psi(t_1))$:

Գրենք U, M_0, M_1 բազմությունների հենման ֆունկցիաները.

$$C(U, \psi) = \|\psi\|,\ C(M_0, \psi) = 3\pi\psi_1 + \pi\|\psi\|,$$

$$C(M_1, \psi) = \pi\|\psi\| :$$

Մաքսիմումի սկզբունքից ստանում ենք

$$(u(t), \psi(t)) = \|\psi(t)\|,\ \|u(t)\| \leq 1:$$

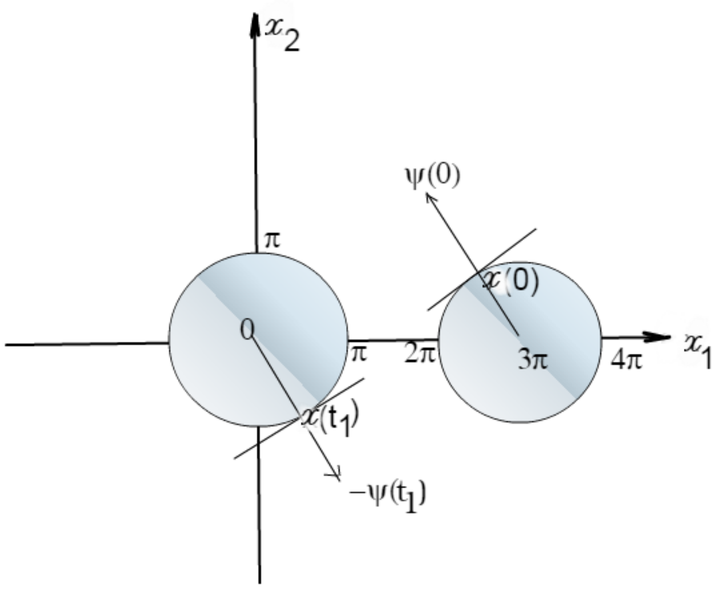
Այսպեղից՝
$$u(t) = \frac{\psi(t)}{\|\psi(t)\|}:$$

Քանի որ $\psi(t) = e^{-tA^T}\psi(0)$, ապա $\|\psi(t)\| = \|\psi(0)\| = 1$: Ուրեմն $u(t) = \psi(t)$: Տրանսվերսալության պայմաններից սպանում ենք

$$x(0) = \begin{pmatrix} 3\pi \\ 0 \end{pmatrix} + \pi\psi(0), \qquad (7.6.13)$$

$$x(t_1) = -\pi\psi(t_1) \qquad (7.6.14)$$

(տես գծ.7.14): Հաշվի առնելով (7.6.13) պայմանը՝ օպտիմալ



Գծ. 7.14: Տրանսվերսալության պայմանները

$x(t)$ հետագծի համար կունենանք հետևյալ տեսքը.

$$x(t) = e^{tA}x(0) + \int\limits_0^t e^{-sA}u(s)ds =$$

$$= e^{tA}[\begin{pmatrix} 3\pi \\ 0 \end{pmatrix} + \pi\psi(0) + \int\limits_0^t e^{-sA}e^{-sA^T}\psi(0)ds] =$$

$$= e^{tA}[\begin{pmatrix} 3\pi \\ 0 \end{pmatrix} + (\pi + t)\psi(0)]:$$

Տեղադրելով այստեղ $t = t_1$՝ (7.6.14) պայմանը կարող ենք գրել հետևյալ տեսքով.

$$e^{t_1A}[\begin{pmatrix} 3\pi \\ 0 \end{pmatrix} + (\pi + t_1)\psi(0)] = -\pi e^{-t_1A^T}\psi(0):$$

Այս հավասարության երկու մասերը բազմապատկելով $e^{t_1A^T}$ մատրիցով՝ կստանանք

$$\begin{pmatrix} 3\pi \\ 0 \end{pmatrix} = -(2\pi + t_1)\psi(0) : \qquad (7.6.15)$$

Վերցնելով այս հավասարության երկու կողմերի մոդուլները՝ կստանանք

$$3\pi = 2\pi + t_1 \Rightarrow t_1 = \pi:$$

Հետևաբար $\psi(0) = (-1, 0)$: Այստեղից՝

$$\psi(t_1) = e^{-t_1A^T}\psi(0) = e^{-\pi A^T}\psi(0) = -E\psi(0) = (1, 0):$$

Այսպիսով, օպտիմալ հետագիծը ունի հետևյալ տեսքը.

$$x(t) = e^{tA}[\begin{pmatrix} 3\pi \\ 0 \end{pmatrix} + (\pi + t)\begin{pmatrix} -1 \\ 0 \end{pmatrix}] =$$

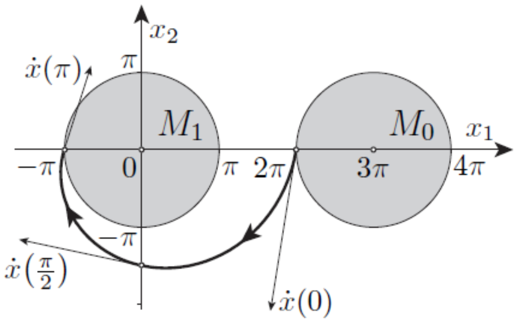
$$= (2\pi - t)\begin{pmatrix} cost \\ sint \end{pmatrix}:$$

Իսկ օպտիմալ կառավարումը կունենա հետևյալ տեսքը.

$$u(t) = \psi(t) = e^{-tA^T}\psi(0) =$$

$$= \begin{pmatrix} cost & sint \\ -sint & cost \end{pmatrix}\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -cost \\ sint \end{pmatrix}:$$

Գծ.7.15-ում պատկերված է $x(t)$ օպտիմալ հետագիծը:



Գծ. 7.15: Օպտիմալ հետագիծը

ԽՆԴԻՐՆԵՐ

1. Լուծել օպտիմալ արագագործության խնդիրը կառավարվող օբյեկտի համար, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով.

$$\begin{cases} \frac{dx_1}{dt} = x_2, \\ \frac{dx_2}{dt} = -x_1 + u, \end{cases}$$

$$\mid u(t) \mid \leq 1, \; t \in [0, T], \; x(0) = (1, 0), \; x(T) = (0, 0):$$

2. Լուծել օպտիմալ արագագործության խնդիրը կառավարվող օբյեկտի համար, որի շարժումը նկարագրվում է դիֆերենցիալ հավասարումների հետևյալ համակարգով.

$$\begin{cases} \frac{dx_1}{dt} = x_2 + u_1, \\ \frac{dx_2}{dt} = -x_1 + u_2, \end{cases}$$

$$u = (u_1, u_2) \in B_1((0, 0)), M_0 = \{(0, 0)\},$$
$$M_1 = \{x \in R^2 / x_1 = 2\pi, \; \mid x_2 \mid \leq 1\}:$$

Օգտագործված գրականության ցանկ

[1] **М. Э. Аббасов**, Методы оптимизации. СПБ, Изд. ВВМ, 2014.

[2] **М. Аоки**, Введение в методы оптимизации, М., Наука, 1977.

[3] **В.М. Алексеев, В. М Тихомиров, С. В. Фомин**, Оптимальное управление, Наука, М., 1979.

[4] **В.М. Алексеев, Э.М. Галеев, В.М. Тихомиров**, Сборник задач по оптимизации, Наук, М., 1984.

[5] **И. Л. Акулич**, Математические программирование в примерах и задачах, Высшая школа, М., 1986.

[6] **И. М. Гельфанд, С. В. Фомин,** Вариационное исчисление, М., Наука, 1961.

[7] **Ф. П. Васильев**, Численные методы решения экстремальных задач, Наука, М., 1980.

[8] **Н.Н. Воробьев,** Числа Фибоначчи, М., Наука, 1969.

[9] **В.И. Благодатских,** Введение в оптимальное управление. Линейная теория, Высшая школа, 2001.

[10] **Ф. Гилл, У.Мюррей, М. Райт,** Практическая оптимизация, Пер. с.англ., М., Мир,1985.

[11] **И. В. Гирсамов,** Лекции по математической теории экстремальных задач, М., Изд.-во МГУ, 1970.

[12] **В. Ф. Демьянов, Б.Н. Малоземов,** Введение в минимакс, М., Наука, 1972.

[13] **В. Г. Карманов,** Математическое программирование, М., Наука, 1975.

[14] **Ю. Н. Киселев, С. Н. Аввакумов, М. В. Орлов,** Оптимальное управление. Линейная теория и приложения, М., Изд., МГУ, 2007.

[15] **Э. Б. Ли, Л. Маркус,** Основы теории оптимального управления, Пер. с.англ., М., Наука,1972.

[16] **Н. Н. Моисеев и др,** Методы оптимизации, М., Наука, 1978.

[17] **Е. Поляк,** Численные методы оптимизации, М., Мир, 1974.

[18] **Е.С. Половинкин Е.С., М. В. Балашов** , Элементы выпуклого и сильно выпуклого анализа, физматлит, М., 2004.

[19] **Л.С. Понтрягин и др.,** Математическая теория оптимальных процессов, М., Наука, 1976.

[20] **А. В. Пантелеев, Т.А. Летова,** Методы оптимизации в примерах и задачах, Высшая школа, М., 2002.

[21] **Б.Н.Пшеничный, Ю. М. Данилин,** Численные методы в экстремальных задачах. Наука, М., 1975.

[22] **Б.Н.Пшеничный**, Выпуклый анализ и экстремальные задачи, Наука, М., 1980.

[23] **Б.Н.Пшеничный,** Метод линеаризации, М., Наука, 1980.

[24] **А. Г. Сухарев, А.В. Тимохов, В.В. Федеров**, Курс методов оптимизации, Наука, М., 1986.

[25] **М.А. Лаврентев,Л.А. Люстерник,** Курс вариационного исчисления, М.: Л., 1950.

[26] **К. Лейхтвейс**, Выпуклые множества, Наука, М., 1985.

[27] **В. В. Воеводин** Линейная алгебра, Наука, М., 1987.

[28] **А. Фиакко, Г. Мак-Кормик,** Нелинейное программирование. Методы последовательной безусловной минимизации, М., Мир, 1972.

[29] **Д. Химельблау,** Прикладное нелинейное программирование, М., Мир, 1975.

[30] **R.T. Rockafellar , J.B. Roger,** Variational Analisis, Springer-Verlag, Berlin, Heidelberg, 1998.

[31] **Վ. Ավետիսյան, Մ. Պողոսյան,** Վարիացիոն հաշիվ և օպտիմալ կառավարում (դասախոսություններ), ԵՊՀ, Երևան, 2008:

[32] **Վ. Ռ. Բարսեղյան**, Վարիացիոն հաշիվ, Ուսումնական ձեռնարկ, ԵՊՀ, Երևան, 2011:

[33] **Ռ. Ա. Խաչատրյան,** Օպտիմիզացիայի մեթոդներ, Ուսումնական ձեռնարկ, ԵՊՀ, Երևան, 2014:

[34] **Ռ.Ավետիսյան, Գ. Տոնոյան, Ս. Սարգսյան, Ա. Պետրոսյան**, Գործույթների հետազոտման մաթեմատիկական մեթոդներ, Ուսումնամեթոդական ձեռնարկ, ԵՊՀ, Երևան, 2017:

ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

ՌԱՖԻԿ ԱՂԱՍՈՒ ԽԱՉԱՏՐՅԱՆ

ՕՊՏԻՄԻԶԱՑԻԱՅԻ ՄԵԹՈԴՆԵՐ

Համակարգչային ձևավորումը՝ Ռ. Խաչատրյանի
Կազմի ձևավորումը՝ Ա. Պատվականյանի
Հրատ. սրբագրումը՝ Լ. Հովհաննիսյանի

Տպագրված է «ՎԱՌՄ» ՍՊԸ-ում:
Ք. Երևան, Տիգրան Մեծի 48, բն. 43

Ստորագրված է տպագրության՝ 20.11.2020:
Չափսը՝ 60x84 1/16: Տպ. մամուլը՝ 19.25:
Տպաքանակը՝ 100:

ԵՊՀ հրատարակչություն
ք. Երևան, 0025, Ալեք Մանուկյան 1
www.publishing.am

ՀՐԱՏԱՐԱԿՉՈՒԹՅՈՒՆ
ԵՐԵՎԱՆ 2020
publishing.ysu.am