

LATVIJAS UNIVERSITĀTE  
FIZIKAS, MATEMĀTIKAS UN OPTOMETRIJAS FAKULTĀTE MATEMĀTIKAS  
NODAĻA

**NEDIFERENCĒJAMU FUNKCIJU  
OPTIMIZĀCIJAS PROBLĒMAS**

BAKALaura DARBS

AUTORS: Kristaps Feikners  
STUDENTA APLIECĪBAS NR.: kf17014  
DARBA VADĪTĀJA: profesore Dr. mat. Inese Bula

RĪGA 2021

## Anotācija

Bakalaura darba nolūks ir izpētīt optimizācijas metodes nediferencējamām funkcijām, lai veiksmīgi spētu tās pielietot dažādu problēmu risināšanā. Lielākoties darbs tiek balstīts uz trīs optimizācijas metodēm, tas ir, gradienta metodi, saišķa metodi un metodi, kuras pamatā nediferencējamu funkciju aizstāj ar diferencējamu. Darba gaitā apskatītās risināšanas tehnikas tiks aprakstītas gan teorētiski, gan arī parādīti praktiski piemēri ar programmā *Wolfram Mathematica* veidotiem grafikiem. Ir aprakstītas metožu pielietošanas iespējas un arī to trūkumi. Bakalaura darbā tiks iepazīstināts ar optimizācijas metožu vēsturi, minēti to pielietojumi dažādās profesionālajās sfērās, kā arī aprakstīti optimizācijas metožu pamatjēdzieni. Pēc paveiktā darba tiks doti ieteikumi metožu izvēlē, kā arī tiks ieskicēti iespējamie darba turpinājuma virzieni. Iegūtie secinājumi ļauj labāk izprast nediferencējamu funkciju optimizācijas metožu pielietošanas iespējas un to īpatnības. Bakalaura darbs ir metodiska rakstura, to var izmantot optimizācijas metožu kursā.

**Atslēgas vārdi:** nediferencējama funkcija, funkciju optimizācija, gradienta metode, saišķa metode, nediferencējamu funkciju aizstāšana ar diferencējamu.

## Abstract

The purpose of the bachelor's thesis is to study optimization methods for non-differentiable functions in order to successfully apply them in solving various problems. The work is mostly based on three optimization methods, that is, the gradient method, the bundle method, and the method based on replacing a non-differentiable function with a differentiable one. The solution techniques discussed in the work will be described theoretically and practically with the examples containing graphs created in the program *Wolfram Mathematica*. Advantages and disadvantages of the methods' applications are described. The bachelor's thesis will introduce the history of optimization methods, mention their applications in various professional fields, as well as describe the basic concepts of optimization methods. At the end of the work, recommendations for the choice of methods will be given, as well as possible directions for further work will be outlined. The obtained conclusions allow to better understand the possibilities of application of non-differentiable function optimization methods and their peculiarities. The bachelor's thesis has a methodological nature, it can be used in the optimization methods' course.

**Keywords:** non-differentiable function, function optimization, gradient method, bundle method, replacement of non-differentiable function with differentiable.

## Saturs

|   |           |
|---|-----------|
| Apzīmējumi . . . . .  | 5         |
| Ievads . . . . .  | 6         |
| <b>1. Pamatjēdzieni un galvenie rezultāti optimizācijas metodēs . . . . .</b>     | <b>7</b>  |
| 1.1. Optimizācijas metožu vēsture . . . . .                                       | 7         |
| 1.2. Optimizācijas metožu pielietojumi . . . . .                                  | 8         |
| 1.3. Optimizāciju metožu pamatjēdzieni . . . . .                                  | 9         |
| 1.4. Apakšgradienta jēdziens . . . . .  | 11        |
| <b>2. Nediferencējamu funkciju optimizācijas piemēri . . . . .</b>                | <b>17</b> |
| <b>3. Risināšanas pamatidejas . . . . .</b>                                       | <b>23</b> |
| 3.1. Problēmas ar gradienta metodi . . . . .                                      | 23        |
| 3.2. Nediferencējamās funkcijas aproksimēšana ar diferencējamu funkciju . . . . . | 29        |
| 3.3. Saišķa ( <i>bundle</i> ) metodes pamatideja . . . . .                        | 32        |
| Secinājumi . . . . .  | 39        |
| Izmantotā literatūra un avoti . . . . .   | 40        |

## Apzīmējumi

$\mathbb{R}$  - reālo skaitļu kopa;

$\mathbb{R}^n$  - n-dimensionālo reālo skaitļu kopa;

$x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  - n-dimensionāls vektors telpā  $\mathbb{R}^n$ ;

$f'(x) = (f'_{x_1}(x), f'_{x_2}(x), \dots, f'_{x_n}(x)) = \text{grad } f$  - funkcijas  $f$  gradients (parciālo atvasinājumu vektors), tiek saukts arī par funkcijas  $f$  pirmo atvasinājumu;

$$f''(x) = \begin{pmatrix} f''_{x_1x_1}(x) & f''_{x_1x_2}(x) & \dots & f''_{x_1x_n}(x) \\ f''_{x_2x_1}(x) & f''_{x_2x_2}(x) & \dots & f''_{x_2x_n}(x) \\ \dots & \dots & \dots & \dots \\ f''_{x_nx_1}(x) & f''_{x_nx_2}(x) & \dots & f''_{x_nx_n}(x) \end{pmatrix} - \text{funkcijas } f \text{ otrais atvasinājums};$$

$(x, y)^T$  - transponētais vektors, t.i.,  $(x, y)^T = \begin{pmatrix} x \\ y \end{pmatrix}$ ;

$|x|$  - reāla skaitļa  $x$  modulis jeb absolūtā vērtība;

$\|x\| = \sqrt{x^T x} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$  - vektora  $x$  Eiklīda norms;

$\varepsilon$  - mazs pozitīvs reāls skaitlis;

■ - piemēru beigu simbols;

LP - lineārā programmēšana;

LPP - lineārās programmēšanas problēma;

NLP - nelineārā programmēšana;

NLPP - nelineārās programmēšanas problēma.

## Ievads

Nediferencējamu funkciju optimizācija ir optimizācijas metožu joma, kas risina optimizācijas problēmas, kurās dažādu iemeslu dēļ nav diferencējamās funkcijas. Funkcijas šajā optimizācijas klasē parasti ir nepārtrauktas, bet tās nav gludas, un bieži satur asus punktus vai stūrus, kas neļauj risināt problēmas ar diferencēšanas palīdzību. Praksē nediferencējamu funkciju optimizācija aptver daudz dažādu problēmu, un to atrisinājumus nevar atrast tikai ar vienu vienīgu metodi. Šādas funkcijas bieži rodas reālās dzīves situācijās kā, piemēram, ekonomikas jomā, kur izmaksu funkcijas parasti ietver asus punktus. Mūsdienās nediferencējamās optimizācijas sfēra turpina attīstīties, jo arvien jaunām problēmām tiek meklēti arvien efektīvāki risinājumi.

Bakalaura darba mērķis ir aplūkot dažādas nediferencējamu funkciju optimizāciju metodes, izprast to pielietojumus, kā arī noteikt metožu trūkumus. Darba uzdevums ir izpētīt sīkāk informācijas avotu piedāvātos piemērus un pa soļiem izveidot pārskatāmu risinājumu shēmu.

Bakalaura darbs sastāv no teorētiskās un praktiskās daļas apvienojuma, kā arī no secinājumiem. Darbā tiek aplūkota optimizācijas metožu vēsture, pielietojumi reālajā dzīvē, kā arī tiek minēti dažādi pamatjēdzieni saistībā ar darba tēmu. Praktiskajā daļā pa soļiem tiek parādīta risināšanas gaita dažādiem nediferencējamu funkciju optimizāciju metožu piemēriem un aplūkoti šo metožu trūkumi.

Bakalaura darbs ir izstrādāts, pamatojoties uz studiju un prakses laikā iegūtajām zināšanām, kā arī dažādiem informācijas un literatūras avotiem.

# 1. Pamatjēdzieni un galvenie rezultāti optimizācijas metodēs

## 1.1. Optimizācijas metožu vēsture

Vispārīgā gadījumā ar jēdzienu "optimizācijas metodes" saprot viena vai vairāku argumentu funkcijas maksimālās vai minimālās vērtības noteikšanu un ekstrēma punktu atrašanu. Pirmie algoritmi ar kuru palīdzību atrod iterāciju ceļā optimizācijas problēmas atrisinājumu, saistīti ar lineāro programmēšanu (LP). Lineārajā programmēšanā vairāku argumentu funkcija ir lineārā formā un mainīgie lielumi apmierina nosacījumu sistēmu, ko veido lineāras vienādības un nevienādības. Nelineārās programmēšanas (NLP) gadījumā optimizējamā funkcija un/vai ierobežojumi nav lineāri ([3]).

Pirmo problēmu lineārajā programmēšanā izveidoja krievu matemātiķis L. V. Kantorovičs (*Leonid Vitaliyevich Kantorovich*, 19.01.1912.-07.04.1986.) un amerikāņu ekonomists F. L. Hičkoks (*Frank Lauren Hitchcock*, 06.03.1875.-31.05.1957.). Viņi darbojās ap transportēšanas problēmu, kas veido LP atzaru. Lai gan franču matemātiķim Žanam - Baptistam Dž. Fornjē (*Jean-Baptiste Joseph Fourier*, 21.03.1768.-16.05.1830.) izskatījās, ka jau ap 1823. gadu bija nojausma par šī priekšmeta potenciālu, tomēr L. V. Kantorovičs 1939. gadā publicēja plašu matemātisko metožu monogrāfiju ražošanas organizēšanā un plānošanā un tiek uzskatīts kā pirmais, kas ievēroja definētas matemātiskas struktūras dažām svarīgām plaša mēroga plānošanas problēmu klasēm. Iespējams, pati pirmā matemātiskās programmēšanas problēma bija optimāla ierobežoto resursu sadale, kuru konstatēja ekonomisti ap 1930. gadu. Pēc Otrā pasaules kara 1947. gadā Amerikas Savienoto Valstu gaisa spēku optimālo programmu zinātniskās aprēķināšanas komanda "Scoop" sāka veikt intensīvu izpēti saistībā ar kādu optimālu resursu sadales problēmu. Šī izpēte vēlāk vainagojās pie slavenās Dž. B. Danciga (*George Bernard Dantzig*, 08.11.1914.-13.05.2005.) LPP atrisināšanas simpleksa algoritma izveides.

H. V. Kūns (*Harold William Kuhn*, 29.07.1925.-02.07.2014.) un A. V. Takers (*Albert William Tucker*, 28.11.1905.-25.01.1995.) 1951. gadā izveidoja nepieciešamos nosacījumus, kuri vēlāk arī kļuva par pietiekamajiem nosacījumiem, lai tie tiktu apmierināti ar NLPP optimālo risinājumu. Šie nosacījumi, kuri ir zināmi kā K-T nosacījumi, lika pamatu daudziem pētījumiem un attīstībai NLP atrisinājumu meklēšanas metodēs. Līdz pat šodienai vēl nav izveidota viena vienīga metode, kas var atrast optimālo risinājumu katrai NLPP, kā tas ir LP metodēs ar simpleks metodi. Pagaidām ir pieejamas dažādas risināšanas metodes specifiska tipa NLPP. Viena no visspēcīgākajām metodēm ir nelineāras problēmas transformācija tā, lai varētu pielietot LPP simpleksa metodi. Starp citām NLPP risināšanas metodēm ir arī gradienta un gradienta projekcijas metodes. Tāpat kā LPP simpleksa metodei, tām tiek izmantotas iteratīvas procedūras, kur katrā solī dodās no viena iespējamā risinājuma uz otru tā, lai mērķa funkcijas vērtība uzlabotos.

Optimizācijas metožu apakšnozares ir izliektā programmēšana, veselu skaitļu programmēšana, stohastiskā programmēšana, dinamiskās programmēšana un citas. Opti-

mizācijas metodēs apakšnozares ir tik daudz, ka visas netika nosauktas un visticamāk nākotnē parādīsies arī jaunas šīs sfēras apakšnozares ([6]).

## 1.2. Optimizācijas metožu pielietojumi

Agrīnie matemātiskās programmēšanas pielietojumi bija saistīti ar militāro plānošanu un dažādu projektu koordinēšanu, kā arī ar ierobežotu resursu efektīvu izmantošanu. Pēdējo piecdesmit gadu laikā matemātiskās programmēšanas metodes tiek veiksmīgi pielietotas gandrīz visās cilvēka darbības jomās, piemēram, rūpniecībā, lauksaimniecībā, inženierzinātnēs, zinātniskajos pētījumos utt.

Nozīmīgi matemātiskās programmēšanas risināšanas metožu pielietojumi ir redzami dažādās statistikas problēmās kā, piemēram, konkrēta eksperimenta izstrādē, vai liela mēroga datu vākšanas apsekojuma plānošanā, vai stohastiskā modeļa izvēlē novēroto datu raksturošanai, vai secinājumu izdarīšanā no pieejamajiem datiem, tas ir, novērtēšanā, hipotēzes pārbaudē un lēmumu pieņemšanā. Lai to visu realizētu, ir nepieciešams izvēlieties mērķa funkciju un minimizēt vai maksimizēt to, ņemot vērā noteiktos ierobežojumus vai nezināmos parametrus un ievadītos datus, piemēram, iesaistītās izmaksas. Klasiskās optimizācijas metodes, kuru pamatā ir diferenciāļi, ir pārāk ierobežojošas un vai nu nav piemērotas, vai arī ir grūti izmantojamas daudzās situācijās, kas rodas statistikas darbā. Iepriekš minētais kopā ar piemērotu skaitlisku algoritmu trūkumu optimizējošu vienādojumu risināšanā ir izvirzījis nopietnus ierobežojumus mērķa funkciju un ierobežojumu izvēlei, kas noved pie dažu neefektīvu statistikas procedūru izstrādes un izmantošanas. Tāpēc pēdējo trīsdesmit gadu laikā ir mēģināts atrast citas optimizācijas metodes, kuras ir plašāk pielietojamas un kuras var ērti izmantot ar pieejamo datoru skaitļošanas jaudu. Viena no šādām metodēm, kurai ir potenciāls palielināt efektīvas statistikas metodoloģijas piemērošanas iespējas, ir matemātiskā programmēšana.

1957. gada pētījumā A. Čārness (*Abraham Charnes*, 04.09.1917.-19.12.1992.) un V. V. Kūpers (*William Wager Cooper*, 23.07.1914.-20.06.2012.) iepazīstināja ar matemātiskās programmēšanas metožu izmantošanu statistikā. Lineārajā regresijā viņi izvēlējās pieeju minimizēt absolūtās novirzes summu, kas ir alternatīva mazākajai kvadrāta metodei, un absolūtās novirzes summas minimizācijas problēmu formulēja kā LPP. 1973. gadā tika ieviesta simpleksa metodes efektīva modifikācija, kas palielināja iespēju izmantot absolūtās novirzes summas minimizācijas regresiju kā alternatīvu klasiskajai regresijai. Bez iepriekš minētā daži citi matemātiskās programmēšanas metožu pielietojumi statistiskās analīzes problēmām ir atrodami klasteru analīzē, dizaina konstruēšanā, uzticamības un kvalitātes kontrolē.

Matemātiskās programmēšanas modeļi un metodes tiek tik plaši izmantoti dažādās disciplīnās, kas rodas gandrīz visās zinātnes nozarēs, rūpniecībā, lauksaimniecībā, inženierzinātnēs, vadībā, plānošanā, sociālajās un ekonomiskajās problēmās, medicīnas zinātnē,

uzņēmējdarbībā, militārajā nozarē, statistikas analizē utt.. Pielietojums ir tik plašs, ka ir grūti sniegt pilnu visu matemātiskās programmēšanas metožu lietojumu sarakstu ([6]).

### 1.3. Optimizāciju metožu pamatjēdzieni

Optimizācijas metodes galvenā ideja ir atrisināt optimizācijas problēmu jeb no dažādām alternatīvu kopām izvēlēties vislabāko elementu, ņemot vērā dotos kritērijus. Dažāda veida optimizācijas problēmas rodas visās kvantitatīvajās disciplīnās, sākot no datorzinātnēm un inženierzinātnēm līdz procesu izpētei un ekonomikai, un matemātikā risinājumu izstrāde ir bijusi aktuāla gadsimtiem ilgi. Vienkāršākajā gadījumā optimizācijas problēma sastāv no reālās funkcijas maksimizēšanas vai minimizēšanas, sistemātiski izvēloties ievades vērtības no atļautās kopas un aprēķinot funkcijas vērtību.

Optimizācijas problēmu var formulēt šādi. Pieņemsim, ka ir dota funkcija  $f : A \rightarrow \mathbb{R}$ , kur  $A \subset \mathbb{R}^n$ . ir jāatrod tāds elements  $x_0 \in A$ , ka  $f(x_0) \leq f(x)$  visiem  $x \in A$  (minimizācijas problēmas gadījumā) vai  $f(x_0) \geq f(x)$  visiem  $x \in A$  (maksimizācijas problēmas gadījumā). Kopu  $A$  arī var aprakstīt ar dažādiem ierobežojumiem (nevienādībām, vienādībām, intervāliem). Dažkārt šādi formulēto problēmu sauc arī par matemātiskās programmēšanas problēmu. Daudzas reālās un teorētiskās problēmas var tikt modelētas pēc šādas vispārīgas struktūras ([4]).

**Piemērs 1.3.1.** Metereoloģiskā stacija nosaka gaisa temperatūru atšķirīgos augstumos (metros virs jūras līmeņa). Visi mērījumi tiek veikti vienā un tajā pašā dienā un laikā, iegūstot datu tabulu ar  $x_i$  (augstums metros virs jūras līmeņa) un  $z_i$  (temperatūra grādos pēc Celsija),  $i = 1, \dots, n$ , kur  $n$  - novērojumu skaits. Novērotās vērtības liecina, ka temperatūra  $z$  samazinās lineārā atkarībā no augstuma  $x$ . Uzdevums ir noskaidrot, kā šī lineārā funkcija  $z = ax + b$  konkrēti izskatās. Problēmu varētu atrisināt, minimizējot attālumus no reālajiem datiem  $z_i$  līdz taisnes punktiem  $ax_i + b$ , t. i.,

$$\sum_{i=1}^n |z_i - (ax_i + b)| \rightarrow \min.$$

Uzdevums satur moduļa funkciju (ja to nelieto, tad savstarpēji izteiksmes  $z_i - (ax_i + b)$  sasummējot, var iegūt 0, kaut novirzes bijušas lielas), kas nav diferencējama. Viens no variantiem ir lietot mazāko kvadrātu metodi, lai noteiktu parametrus  $a$  un  $b$ . Šajā nolūkā ir jārisina minimizācijas problēma:

$$\sum_{i=1}^n (z_i - (ax_i + b))^2 \rightarrow \min. \blacksquare$$

Parasti  $A$  ir kādas Eiklīda telpas  $\mathbb{R}^n$  apakškopa, kuru visbiežāk nosaka ar ierobežojumiem, vienādību vai nevienādību formā, kas kopas  $A$  elementiem ir jāapmierina. Funkciju

$f$  dēvē dažādi, piemēram, minimizācijas problēmās to dēvē par mērķa funkciju, zaudējumu funkciju vai izmaksu funkciju, bet maksimizācijas uzdevumos to sauc par lietderības funkciju vai atbilstības funkciju vai noteiktās sfērās par enerģijas funkciju. Pieļaujama atrisinājums, kas minimizē vai maksimizē mērķa funkciju, tiek saukts par optimālo atrisinājumu ([7]).

**Definīcija 1.3.2.** ([1]) Pieņemsim, ka punkts  $x_0$  pieder funkcijas  $f$  definīcijas apgabalam  $A$ .

1.  $f(x_0)$  sauc par funkcijas  $f$  lielāko vērtību kopā  $A$ , ja

$$\forall x \in A \quad f(x_0) \geq f(x).$$

2.  $f(x_0)$  sauc par funkcijas  $f$  mazāko vērtību kopā  $A$ , ja

$$\forall x \in A \quad f(x_0) \leq f(x).$$

Funkcijas lielāko un mazāko vērtību kopā  $A$  sauc par ekstremālajām vērtībām. Punktus, kuros tiek sasniegtas ekstremālās vērtības, sauc par ekstrēmu punktiem. Punktu  $x_0$  1. gadījumā sauc par maksimuma punktu, bet 2. gadījumā par minimuma punktu. Vienargumenta funkcijas gadījumā parasti ar minimuma vai maksimuma punktu saprot  $(x_0, f(x_0))$ . Definīcijā 1.3.2. minētos ekstrēmus mēdz saukt par globāliem ekstrēmiem.

**Definīcija 1.3.3.** ([1]) Pieņemsim, ka punkts  $x_0$  pieder funkcijas  $f$  definīcijas apgabalam  $A$ .

1.  $f(x_0)$  sauc par funkcijas  $f$  lokālo maksimālo vērtību, ja eksistē tāda punkta  $x_0$  apkārtnē  $U(x_0)$ , ka  $f(x_0)$  ir globālā maksimālā vērtība kopā  $U(x_0) \cap A$ , t.i.,

$$\forall x \in U(x_0) \cap A \quad f(x_0) \geq f(x).$$

2.  $f(x_0)$  sauc par funkcijas  $f$  lokālo minimālo vērtību, ja eksistē tāda punkta  $x_0$  apkārtnē  $U(x_0)$ , ka  $f(x_0)$  ir globālā minimālā vērtība kopā  $U(x_0) \cap A$ , t.i.,

$$\forall x \in U(x_0) \cap A \quad f(x_0) \leq f(x).$$

$f(x_0)$  sauc par lokālo ekstremālo vērtību, ja tā ir lokālā minimālā vai maksimālā vērtība. Ievērosim, ka jebkura globālā ekstremālā vērtība ir arī lokālā ekstremālā vērtība. Ja Definīcijās 1.3.2 un 1.3.3 ir stringrās nevienādības, tad tiek runāts par stingriem ekstrēma punktiem.

Parasti, ja vien minimizācijas uzdevumā mērķa funkcija nav izliekta, lokālie minimumi var būt vairāki. Izliektas funkcijas problēmā, ja lokālais minimums ir iekšpusē

pieļaujamajam reģionam, nevis uz pieļaujamo elementu kopas robežas, tad tas ir arī globālais minimums, bet neizliktas funkcijas problēmai var būt vairāk nekā viens lokālais minimums, no kuriem visiem nav jābūt globālajiem minimumiem ([5]).

Ja funkcija ir diferencējama, tad ekstrēma eksistences nepieciešamie nosacījumi formulējami šādi.

**Teorēma 1.3.4.** ([1]) Tiek pieņemts, ka  $x_*$  ir funkcijas  $f$  lokālā ekstrēma punkts telpā  $\mathbb{R}^n$  un funkcija  $f$  ir diferencējama šajā punktā. Tad

$$f'(x_*) = 0.$$

Ja funkcija  $f$  ir diferencējama punkta  $x_*$  apkārtne un divreiz diferencējama punktā  $x_*$ , tad lokālā minimuma punktā izpildās  $f''(x_*) \geq 0$  un lokālā maksimuma punktā izpildās  $f''(x_*) \leq 0$ .

Kā ekstrēma eksistences pietiekamos nosacījumus atzīmēsim šādu teorēmu, kurā nepieciešams funkcijas otrās kārtas atvasinājums.

**Teorēma 1.3.5.** ([1]) Tiek pieņemts, ka funkcija  $f$  stacionārā punkta  $x_*$   $\varepsilon$ -apkārtne  $U(x_*, \varepsilon)$  ir diferencējama un divreiz diferencējama punktā  $x_*$ .

Ja  $f''(x_*) > 0$ , tad  $x_*$  ir funkcijas  $f$  stingrs lokāls minimums.

Ja  $f''(x_*) < 0$ , tad  $x_*$  ir funkcijas  $f$  stingrs lokāls maksimums.

Ja kvadrātiskā forma  $\langle f''(x_*)h, h \rangle$  pieņem vērtības ar dažādām zīmēm, tad funkcijai  $f$  punktā  $x_*$  nav ekstrēms.

Ļoti nozīmīgs rezultāts ekstrēmu eksistences noteikšanā ir Veierštrāsa teorēma.

**Teorēma 1.3.6.** ([1]) Pieņemsim, ka  $A \subset \mathbb{R}^n$  ir slēgta un ierobežota kopa un ka  $f : A \rightarrow \mathbb{R}$  ir nepārtraukta funkcija. Tad funkcijai  $f$  kopā  $A$  eksistē gan minimuma punkts, gan maksimuma punkts, t.i.,

$$\exists x^*, x^{**} \in A \quad \forall x \in A \quad f(x^*) \leq f(x), \quad f(x^{**}) \geq f(x).$$

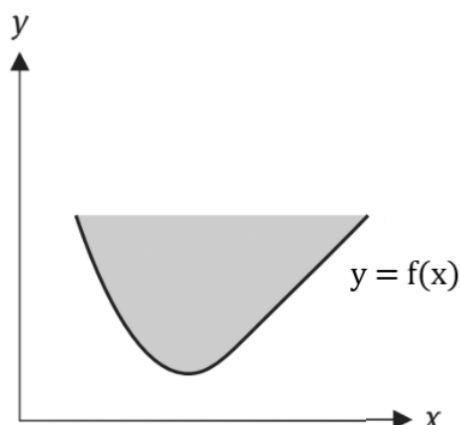
## 1.4. Apakšgradienta jēdziens

Izveidosim apakšgradienta jēdzienu, kas vispārina izliktas funkcijas  $f$  gradienta jēdzienu punktā, kuros funkcija  $f$  nav diferencējama. Šajā nolūkā vispirms definēsim epigrafa un izliktas kopas atbalsta hiperplakni.

**Definīcija 1.4.1.** ([7]) Pieņemsim, ka dota funkcija  $f : M \rightarrow \mathbb{R}$ ,  $M \subset \mathbb{R}^n$ . Par funkcijas  $f$  epigrafu sauc šādu kopu

$$epi(f) = \{(x, r) \in \mathbb{R}^{n+1} \mid x \in M, r \geq f(x)\}.$$

$n=1$  un  $n=2$  gadījumā epigrafs ir visu to kopas  $M$  punktu kopa, kas atrodas virs funkcijas  $f$  grafika (skatīt 1.4.1. att.). Acīmredzami, ka funkcija ir izliekta tad un tikai tad, ja tās



1.4.1. att. Funkcijas  $y=f(x)$  epigrafs.

epigrafs ir izliekta kopa.

**Definīcija 1.4.2.** ([7]) Pieņemsim, ka dota kopa  $M \subset \mathbb{R}^n$  un hiperplakne

$$H = \{x \in \mathbb{R}^n \mid a^T x = b\}, \quad a \in \mathbb{R}^n, \quad b \in \mathbb{R}.$$

Par hiperplaknes  $H$  veidotām puslīnijām sauc kopas

$$S^+ = \{x \in \mathbb{R}^n \mid a^T x \geq b\} \quad \text{un} \quad S^- = \{x \in \mathbb{R}^n \mid a^T x \leq b\}.$$

Hiperplakni  $H$  sauc par kopas  $M$  atbalsta hiperplakni punktā  $x_0$ , ja  $x_0 \in H \cap M$  un  $M$  pilnībā ietilpst vienā no puslīnijām  $S^+$  vai  $S^-$ .

**Piemērs 1.4.3.**

Pieņemsim, ka dots vienības riņķis plaknē

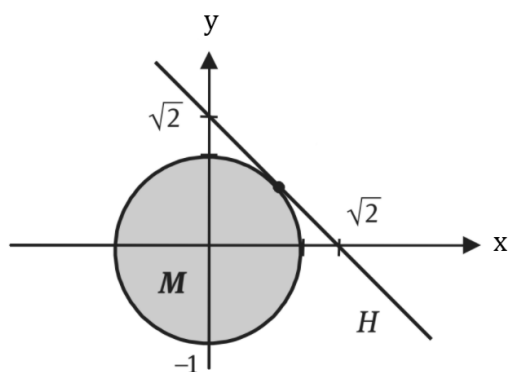
$$M = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}.$$

Šīs kopas atbalsta hiperplakne punktā  $x^0 = (\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}})$  ir taisne (skatīt 1.4.2. att.)

$$M = \{(x, y) \in \mathbb{R}^2 \mid x + y = \sqrt{2}\}.$$

Bet izvēloties citu punktu, piemēram,  $x^0 = (-1; 0)$  atbalsta hiperplakne būs cita taisne

$$M = \{(x, y) \in \mathbb{R}^2 \mid x = -1\}.$$

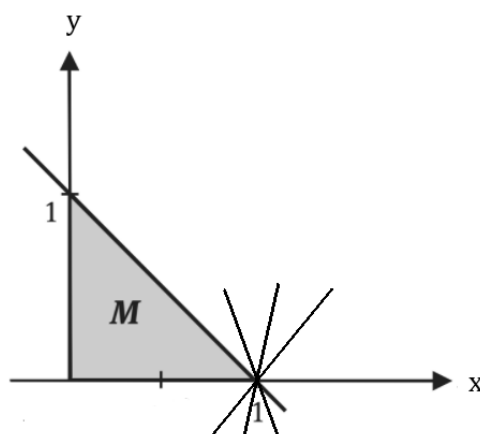


1.4.2. att. Riņķim punktā  $(\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}})$  novilkta atbalsta hiperplakne (taisne).

Kopai  $M$  var būt vienā punktā arī vairākas atbalsta hiperplaknes. Piemēram, pieņemsim, ka dota kopa

$$M = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq 1, x \geq 0, y \geq 0\}$$

(skatīt 1.4.3. att.). Punktā  $(1; 0)$  var novilkt daudzas atbalsta hiperplaknes - šajā gadījumā taisnes (skatīt 1.4.3. att.). Bet piemēram, punktā  $(0; \frac{1}{2})$  var novilkt tikai vienu



1.4.3. att. Trijstūrim punktā  $(1; 0)$  novilkta atbalsta hiperplaknes (taisnes).

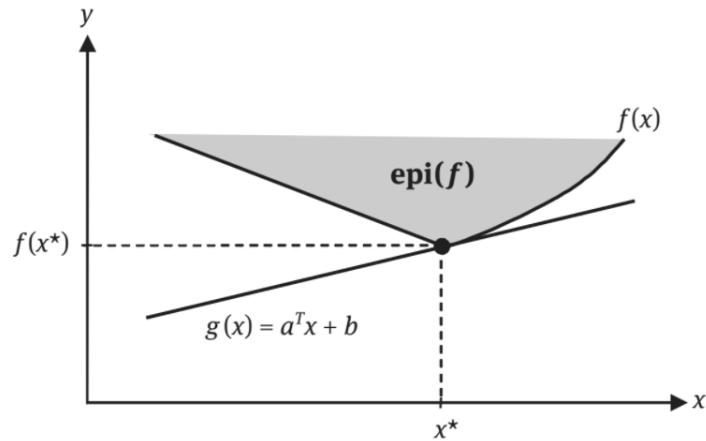
$x = 0$ , pie tam šī atbalsta hiperplakne būs viena un tā pati visiem punktiem  $(0; \alpha)$ , kur  $\alpha \in (0; 1)$ . ■

**Definīcija 1.4.4.** ([7]) Pieņemsim, ka  $f : M \rightarrow \mathbb{R}$  ir izliekta funkcija,  $M \subset \mathbb{R}^n$  - izliekta kopa. Funkciju  $g(x) = a^T x + b$  sauc par funkcijas  $f$  atbalsta funkciju punktā  $x^* \in M$ , ja funkcijas  $g(x)$  grafiks ir funkcijas  $f$  epigrafa atbalsta hiperplakne punktā  $(x^*; f(x^*))$ , t.i., ja

$$f(x^*) = g(x^*) \tag{1.4.1}$$

$$f(x) \geq g(x) \tag{1.4.2}$$

visiem  $x \in M$  (skatīt 1.4.4. att.). Tā kā izpildās (1.4.1), tad atbalsta funkcijas parametrs



1.4.4. att. Funkcijas  $f$  atblasta funkcija ir  $g(x)$ .

$b$  apmierina vienādību  $b = f(x^*) - a^T x^*$ . Tātad izliektas funkcijas  $f$  atbalsta funkcijas punktā  $x^*$  ir funkcijas izskatā

$$g(x) = a^T x + f(x^*) - a^T x^* = f(x^*) + a^T(x - x^*), \quad (1.4.3)$$

kur vektors  $a$  apmierina šādu nevienādību (jo izpildās (1.4.2))

$$\forall x \in M \quad f(x) \geq f(x^*) + a^T(x - x^*). \quad (1.4.4)$$

Ievērosim, ka vektors  $a$  ir funkcijas  $g(x)$  gradients. No šejienes tad radies apakšgradienta jēdziens.

**Definīcija 1.4.5.** ([7]) Pieņemsim, ka  $f : M \rightarrow \mathbb{R}$  ir izliekta funkcija,  $M$ -izliekta kopa. Vektora  $a \in \mathbb{R}^n$  sauc par funkcijas  $f$  apakšgradientu punktā  $x^*$ , ja tas ir funkcijas  $f$  atblasta funkcijas  $g$  gradients punktā  $x^*$ , t.i., ja  $a$  apmierina nosacījumu (1.4.4). Funkcijas  $f$  apakšgradientu kopu punktā  $x^*$  sauc par funkcijas  $f$  apakšdiferenciāli punktā  $x^*$  un apzīmē ar  $\partial f(x^*)$ . Atzīmēsim, ka  $\partial f(x^*) = \{\text{grad} f(x^*)\}$ , ja  $f$  ir diferencējama funkcija punktā  $x^*$ , pretējā gadījumā  $\partial f(x^*)$  ir bezgalīga kopa.

**Piemērs 1.4.6.**

Noskaidrosim apakšdiferenciāli punktā  $x^* = (0; 0)$  izliektai funkcijai  $f(x, y) = |x| + |y|$  (skatīt 1.4.5. att.). Šīs funkcijas grafiks ir virsma (konuss ar virsotni  $(0; 0; 0)$ ), kura šķautnes iet caur punktiem  $(1; 0; 1)$ ,  $(0; 1; 1)$ ,  $(-1; 0; 1)$ ,  $(0; -1; 1)$ . Dotās funkcijas gadījumā punktā  $(0; 0)$  nevienādība (1.4.4) ir pierakstīta kā  $\forall x, y \in \mathbb{R}$

$$|x| + |y| \geq a_1 x + a_2 y,$$

vai ekvivalentā formā

$$|x| \geq a_1 x \text{ un } |y| \geq a_2 y,$$

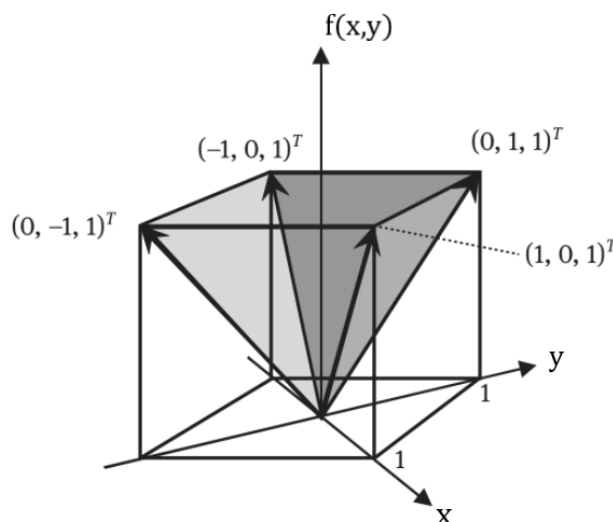
vai ekvivalenti

$$|x| \geq |a_1 x| \text{ un } |y| \geq |a_2 y|$$

visiem  $x, y \in \mathbb{R}$ , iegūsim, ka

$$|a_1| \leq 1 \text{ un } |a_2| \leq 1.$$

Tātad vektors  $a = (a_1, a_2)^T$  ir funkcijas  $f$  apakšgradients punktā  $(0; 0)$  tikai tad, ja  $|a_1| \leq 1$  un  $|a_2| \leq 1$  ([7]). ■



1.4.5. att. Piemēra 1.4.6 funkcijas  $f(x,y)$  grafiks.

Parasti nediferencējamas optimizācijas problēmas ir formā

$$\min_{x \in \mathbb{R}^n} f(x),$$

kur  $f$  ir ar vērtībām no reālas kopas  $\mathbb{R}$ , nepārtraukta un nediferencējama funkcija. Funkcijas  $f$  izliektība nozīmē, ka tai ir vismaz viena atbalstoša hiperplakne katrā  $\mathbb{R}^n$  punktā. Virziena koeficienti šādām hiperplaknēm veido apakšgradientu kopu, kas ir apakšdiferenciāla kopa. Diferencējamos punktos ir viena vienīga atbalstošā hiperplakne, kuras virziena koeficients ir gradients. Nediferencējamos punktos ir bezgalīgi liela apakšgradientu kopa un līdz ar to arī bezgalīgi liela atbalstošo hiperplakņu kopa.

Funkcijas  $f$  atbalstošā hiperplakne punktā  $x_0$  ir dota kā

$$y = f(x_0) + a_0^T(x - x_0),$$

kur  $a_0$  ir jebkurš apakšdiferenciāla  $\partial f(x_0)$  funkcijas  $f$  elements punktā  $x_0$ . Par cik tā ir atbalstošā hiperplakne, tas noved pie apakšgradients nevienādības (1.4.4). Apakšgradianti ir definēti pēc šīs nevienādības.

Noteikt visu apakšdiferenciāļu kopu ir ļoti sarežģīts vai pat neiespējams uzdevums.

Ja funkcija  $f$  ir daudzskaldnis, tad apakšdiferenciāla ekstrēmu punktu skaits var būt eksponenciāls pamatā esošās telpas dimensijā. Pilnīgu apakšdiferenciāla aprakstu var panākt vienkāršai situācijai gadījumā, ja  $f$  ir galīga daudzuma izliektu diferencējamu funkciju maksimums:  $f(x) = \max_{i \in I} f_i(x)$ . Tad apakšdiferenciālis  $\partial f(x_0)$  ir dots kā

$$\partial f(x_0) = \left\{ \sum_{i \in I(x_0)} \alpha_i \operatorname{grad} f_i(x_0) : \sum_{i \in I(x_0)} \alpha_i = 1, \alpha_i \geq 0 \right\},$$

$$I(x_0) = \{i : f_i(x_0) = f(x_0)\}.$$

Nediferencējamā optimizācijā visa apakšdiferenciāļu kopa nekad netiek aprēķināta. Apakšgradientus aprēķina tikai, kad tas ir nepieciešams, un bieži vien pietiek tikai ar vienu elementu ([2]).

## 2. Nediferencējamu funkciju optimizācijas piemēri

Izņemot dažas atrisinājumu meklēšanas tehnikas, lielākajā daļā optimizācijas problēmu tiek pieprasīts, lai funkcijas vismaz vienu reizi būtu nepārtraukti diferencējamas. Tomēr daudzos lietojumos ekonomikā, mehānikā, inženierzinātnēs u.c. parādās funkcijas, kurām nav šī raksturlieluma.

### Piemērs 2.1 (gabaliem lineāra mērķa funkcija).

Uzņēmums, kas vēlas iegūt daudzumu  $a$  konkrētas produkcijas  $P$ , pieprasa produkciju no  $n$  ražotājiem apjomos  $P_1, \dots, P_n$ . Ražotājs  $P_j$  var sagādāt maksimālo daudzumu  $m_j$  un viņa cena par produkciju  $P$  apjomā  $x_j$  ir apzīmēta ar  $f_j(x_j)$  ( $j = 1, \dots, n$ ). Ja uzņēmums vēlas samazināt kopējās izmaksas, tad rodās atdalāma problēma:

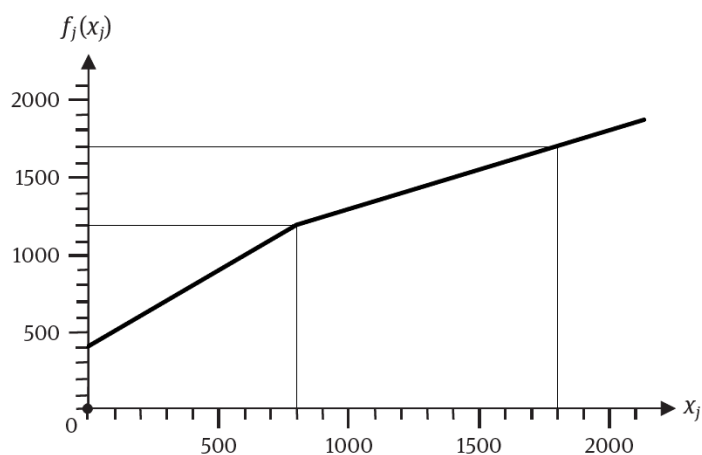
$$\min \sum_{j=1}^n f_j(x_j)$$

$$\sum_{j=1}^n x_j \leq a$$

$$0 \leq x_j \leq m_j, j = 1, \dots, n.$$

**Definīcija 2.2.** ([7]) Funkciju  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sauc par atdalāmu, ja tā sastāv no  $n$  funkciju summas, kur katrai no funkcijām ir tikai viens vienīgs mainīgais, t.i.,  $f(x_1, \dots, x_n) = \sum_{i=1}^n f_i(x_i) = f_1(x_1) + \dots + f_n(x_n)$ .

Piemēram,  $f(x, y) = 3x + y^2$  ir atdalāma funkcija, jo tā sastāv no divu funkciju  $f_1(x) = 3x$  un  $f_2(y) = y^2$  summas.



2.1. att. Piemēra 2.1. j-tā ražotāja izmaksu funkcija.

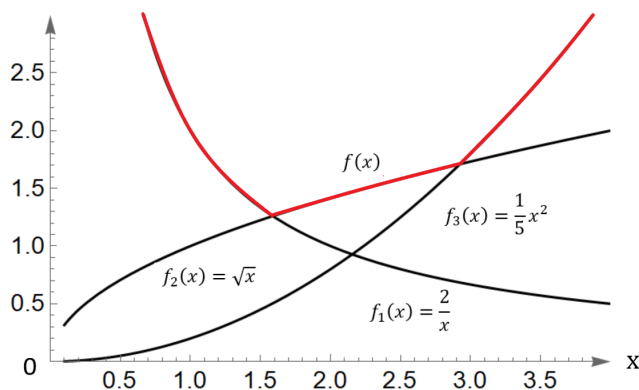
Atsevišķās izmaksas  $f_j(x_j)$  parasti ir augoša, ielikta un gabaliem lineāra funkcija. Kā par piemēru pieņemsim, ka ražošanas ierīces (mašīnas) sagatavošana darbam izmaksā 400 dolāri. Pirmās astoņsimts preces  $P$  vienības tiek pārdotas par 1 dolāru gabalā un par lielākiem apjomiem ražotājs  $P_j$  piešķir 0,5 dolāru lielu atlaidi katrai vienībai. Tiek iegūta funkcija  $f_j(x_j)$  (skatīt 2.1. att.), kura nav diferencējama punktos  $x_j = 0$  un  $x_j = 800$  (ievērosim, ka  $f_j(0) = 0$ ) ([7]). ■

## Piemērs 2.2 (minimaks problēma).

Daudzos pielietojumos minimizējamā funkcija var tikt uzskatīta kā maksimums no funkcijām  $f_1, \dots, f_m: \mathbb{R}^n \rightarrow \mathbb{R}$ , tas ir, problēmai ir šādā formā

$$\min_{x \in \mathbb{R}^n} \max_{i, \dots, m} f_i(x). \quad (2.1)$$

Pat ja visas funkcijas  $f_i$  ir diferencējamas, funkcijai  $f(x) := \max_{i, \dots, m} f_i(x)$  būs "stūri", kur diferencēšana neizdosies. Tas būs divu funkciju  $f_i$  un  $f_j$  krustpunktos (skatīt 2.2. att., kurā dotas funkcijas  $f_1(x) = \frac{2}{x}$ ,  $f_2(x) = \sqrt{x}$ ,  $f_3(x) = \frac{1}{5}x^2$  definētas  $\forall x \in [0, 1; +\infty)$ .



2.2. att. Minimaksa problēmas nediferencējama funkcija.

Iepriekš minētais minimax formulējums satur daudzas apakšproblēmas, piemēram, vektorvērtas funkcijas minimizāciju  $l_1$  normā vai  $l_\infty$  normā, tas ir,

$$\min_{x \in \mathbb{R}^n} \|[f_1(x), \dots, f_m(x)]^T\|_1, \quad (2.2)$$

$$\min_{x \in \mathbb{R}^n} \|[f_1(x), \dots, f_m(x)]^T\|_\infty. \quad (2.3)$$

Pārrakstot formā (2.1), problēmas (2.2) un (2.3) kļūst par

$$\min_{x \in \mathbb{R}^n} \max_{\pm} \{\pm f_1(x), \pm f_2(x), \dots, \pm f_m(x)\}, \quad (2.2.1)$$

$$\min_{x \in \mathbb{R}^n} \max \{f_1(x), -f_1(x), \dots, f_m(x), -f_m(x)\}. \quad (2.3.1)$$

Izteiksmē (2.2.1) ir jāmaksimizē vairāk kā  $2^m$  iespējamās + un - kombinācijas. Vienādojuma (2.1) speciālgadījums ir diskrēts lineārs Čebiševa tuvinājums. Tātad, ja dota "sarežģīta" funkcija  $h : \mathbb{R} \rightarrow \mathbb{R}$ ,  $n$  "vienkāršas" funkcijas  $g_1, \dots, g_n$  un  $m$  punkti  $t_1, \dots, t_m \in \mathbb{R}$  ( $n < m$ ), tad uzdevums ir aproksimēt funkciju  $h$  ar funkciju  $g_j$  lineārajām kombinācijām tā, lai maksimālais attālums starp  $h$  un to tuvinājumu punktos  $t_j$  tiktu minimizēts. Mēs iegūstam problēmu

$$\min_{x \in \mathbb{R}^n} \max_{i, \dots, m} \left| h(t_i) - \sum_{j=1}^n x_j g_j(t_i) \right|, \quad i = 1, \dots, m. \quad (2.4)$$

Apzīmējot  $f_i(x) := h(t_i) - \sum_{j=1}^n x_j g_j(t_i)$ , kur  $i = 1, \dots, m$ , mēs varam pārrakstīt vienādojumu (2.4) kā (2.3.1) ([8]). ■

### Piemērs 2.3.

Lielākajai daļai metožu, lai atrisinātu nosacīto ekstrēmu problēmu

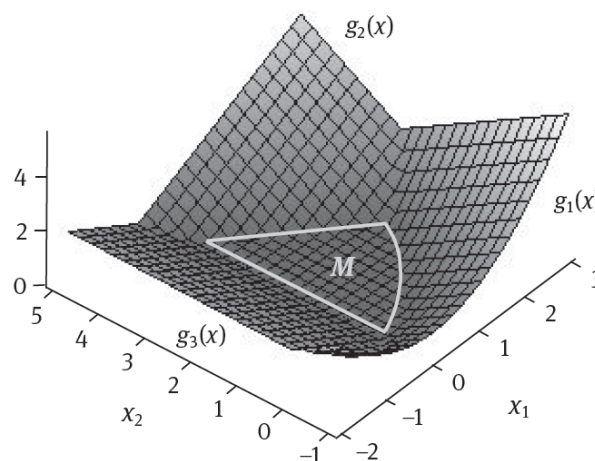
$$\min f(x) \quad (2.5)$$

$$g_i(x) \leq 0, \quad i = 1 \dots m$$

ir nepieciešams pieļaujams punkts, lai sāktu iterāciju procesu. Pieļaujamu punktu  $x^*$  problēmai (2.5) var iegūt, atrisinot problēmu

$$\min_{x \in \mathbb{R}^n} \max \{g_1(x), \dots, g_m(x), 0\}, \quad (2.6)$$

kas ir vēl viens speciālgadījums sākotnējai problēmai (2.5). Protams, funkcija  $h(x) := \max \{g_1(x), \dots, g_m(x), 0\}$  ir nulle tad un tikai tad, ja  $x$  ir pieļaujams punkts problēmai (2.5). Funkcija  $h(x)$  ir ilustrēta 2.3. attēlā gadījumā, ja  $n=2$ ,  $m=3$ ,



2.3. att. Piemēra 2.3 ilustrācija.

$g_1(x) = \frac{1}{2}x_1^2 - x_2$ ,  $g_2(x) = x_1 + x_2 - 4$ ,  $g_3(x) = -x_1$ . Funkcija  $h(x)$  nav diferencējama vairākos "krustpunktos", tas ir, punktos, kur divas no četrām funkcijām  $g_1(x)$ ,  $g_2(x)$ ,  $g_3(x)$ , un  $g_4(x) := 0$  sakrīt. Speciālgadījumā, funkcija  $h(x)$  nav diferencējama problēmas (2.5) pieļaujamā apgabala robežpunktos ([7]). ■

**Dekompozīcija** ([7,8]). Lineāru problēmu var atrisināt divos soļos, uzrakstot to formā

$$\min(c^T x + d^T y) \quad (2.7)$$

$$Ax + By \leq b,$$

kur  $A, B, c, d, b$  ir attiecīgo dimensiju matricas un vektori  $x$  un  $y$  apzīmē attiecīgā vektora mainīgos. Pirmajā solī apakšproblēmai tiek atrasts atrisinājums

$$f(x) := \min_y \{d^T y \mid By \leq b - Ax\} \quad (2.8)$$

pie patvaļīga vektora  $x$ , kurš ir pieļaujams problēmai (2.7). Otrajā solī tiek atrisināta "atlikusī problēma"

$$\min_x \{c^T x + f(x)\}. \quad (2.9)$$

Atrisinājums problēmai (2.7) tagad ir dots kā  $(x^*, y^*)$ , kur  $x^*$  ir atrisinājums problēmai (2.9) un  $y^*$  ir atrisinājums problēmai (2.8), kur  $x = x^*$ . Problēmai (2.7) optimālā vērtība ir  $c^T x^* + f(x^*)$ . Mērķa funkcija  $c^T x + f(x)$  no problēmas (2.9) parasti nav visur diferencējama.

#### Piemērs 2.4.

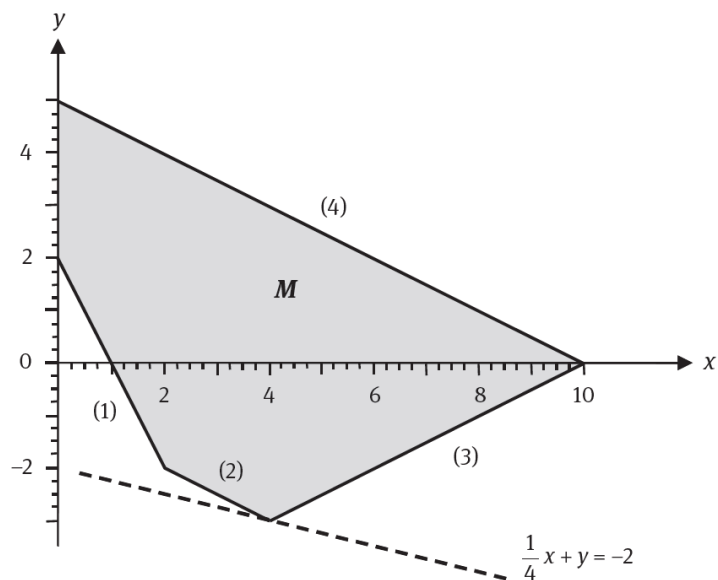
Apskatīsim problēmu

$$\min\left(\frac{1}{4}x + y\right) \quad (2.10)$$

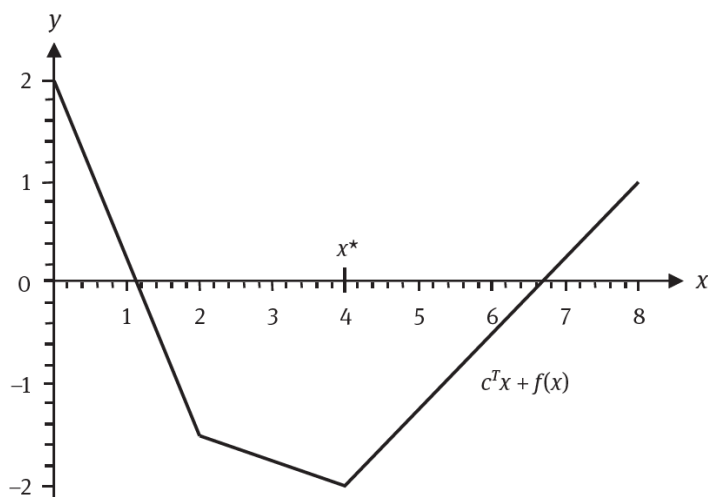
$$\begin{cases} -2x - y \leq -2, & (1) \\ -\frac{1}{2}x - y \leq 1, & (2) \\ \frac{1}{2}x - y \leq 5, & (3) \\ \frac{1}{2}x + y \leq 5, & (4) \\ -x \leq 0. & (5) \end{cases}$$

ar optimālo atrisinājumu  $(4, -3)^T$  (skatīt 2.4. att.). Problēma satur  $c = \frac{1}{4}$ ,  $d = 1$ ,

$$A = \begin{pmatrix} -2 \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -1 \end{pmatrix}, \quad B = \begin{pmatrix} -1 \\ -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 1 \\ 5 \\ 5 \\ 0 \end{pmatrix},$$



2.4. att. Problēmas (2.10) grafiskais atrisinājums.



2.5. att. Problēmas (2.9) mērķa funkcija.

$$\begin{aligned}
 f(x) &:= \min_y \{d^T y \mid B y \leq b - A x\} = \\
 &= \min_y \{y \mid y \geq 2 - 2x, y \geq -1 - \frac{1}{2}x, y \geq -5 + \frac{1}{2}x, y \leq 5 - \frac{1}{2}x\} = \\
 &= \begin{cases} 2 - 2x, & 0 \leq x \leq 2; \\ -1 - \frac{1}{2}x, & 2 < x \leq 4; \\ -5 + \frac{1}{2}x, & 4 < x \leq 10. \end{cases}
 \end{aligned}$$

Jāņem vērā, ka grafiks  $f(x)$  sakrīt ar kopas  $M$  apakšējo robežu (skatīt 2.4. att.), tādēļ

problēmas mērķa funkcija (2.9) ir šāda:

$$c^T x + f(x) = \frac{1}{4}x + f(x) = \begin{cases} 2 - \frac{7}{4}x, & 0 \leq x \leq 2; \\ -1 - \frac{1}{4}x, & 2 < x \leq 4; \\ -5 + \frac{3}{4}x, & 4 < x \leq 10 \end{cases}$$

(skatīt 2.5. att.). Funkcija nav diferencējama punktā  $x = 2$  un minimuma punktā  $x^* = 4$ . Tādējādi problēmas (2.10) optimālais atrisinājums ir  $(x^*, y^*)^T = (4, -3)^T$ , kas saktīt ar grafisko atrisinājumu (skatīt 2.4. att.).

Praksē iepriekš minētā atdalīšana ir noderīga tikai priekš liela izmēra lineārām problēmām, kur matricām  $A$  un  $B$  problēmā (2.7) ir specifiska struktūra, kas atvieglo apakšproblēmas (2.8) atrisinājumu ([7]). ■

Neskaitot iepriekš minētos piemērus, nediferencējama mērķa funkcija var rasties arī citos pielietojumos. Piemēram, soda funkciju metode, kas satur moduli, parasti nav diferencējama visos punktos. Duālajās problēmās mērķa funkcija bieži nav diferencējama. Arī kopējās transporta izmaksas

$$f(x_1, x_2) = s \sum_{i=1}^n q_i \sqrt{(x_1 - a_i)^2 + (x_2 - b_i)^2},$$

kas jāminimizē lokācijas problēmā, nav diferencējamas punktos  $(a_i, b_i)$  ([7]).

### 3. Risināšanas pamatidejas

Iepriekš minētajos piemēros un lielākajā daļā pielietojumu mērķa funkcija ir diferencējama "gandrīz visos" punktos. Šis fakts liek domāt, ka ideja, piemēram, gradientu metode vai kvazi-Nūtona metode, var tikt pielietota, lai minimizētu tāda tipa funkcijas, gaidot, ka algoritms "parasti" ģenerēs diferencējamus punktus. Turklāt, ja iterāciju punkts  $x^k$  nav diferencējams, tad iterāciju metode var tikt atsākta ar "diferencējamu punktu" tuvu  $x^k$ . Tomēr nākošie piemēri parāda, ka šāda procedūra ne vienmēr strādā.

#### 3.1. Problēmas ar gradienta metodi

Tiek pieņemts, ka iterācijas sākumpunkts  $x_0$  ir jau izvēlēts. Gradienta metode paredz izveidot punktu virkni  $(x^k)_{k \in \mathbb{N}}$  pēc sekojoša likuma

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

Skaitli  $\alpha_k$  iepriekš minētajā formulā sauc par soļa garumu vai vienkārši par gradienta metodes soli. Ja  $f'(x)^k \neq 0$ , tad pie pietiekoši maziem  $\alpha_k > 0$  var garantēt dilstošu funkcijas vērtību virkni:

$$f(x^{k+1}) < f(x^k).$$

Viena no populārākajām metodēm, lai noteiktu gradienta metodes soļa lielumu ir ātrākā virziena metode. Uz stara  $x = x^k - \alpha \text{grad } f(x^k)$ ,  $\alpha \geq 0$ , kas iziet no punkta  $x^k$  pretēji gradienta virzienā, tiek izveidota viena argumenta  $\alpha$  funkcija

$$g_k(\alpha) = f(x^k - \alpha \text{grad } f(x^k)), \alpha \geq 0.$$

Solis  $\alpha_k$  tiek definēts kā vienargumenta funkcijas  $g_k(\alpha)$  minimuma punkts.

**Definīcija 3.1.1** ([1]) Kopu  $M \subset \mathbb{R}^n$  sauc par izliektu, ja

$$\forall x \in M \forall y \in M \forall \lambda \in [0; 1] : \lambda x + (1 - \lambda)y \in M.$$

**Definīcija 3.1.2** ([1]) Pieņemsim, ka dota funkcija  $f$ , kas definēta izliektā kopā  $M \subset \mathbb{R}^n$ .

1. Funkciju  $f$  sauc par izliektu uz leju, ja

$$\forall x \in M \forall y \in M \forall \lambda \in (0; 1) : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

2. Funkciju  $f$  sauc par izliektu uz augšu, ja

$$\forall x \in M \forall y \in M \forall \lambda \in (0; 1) : f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y).$$

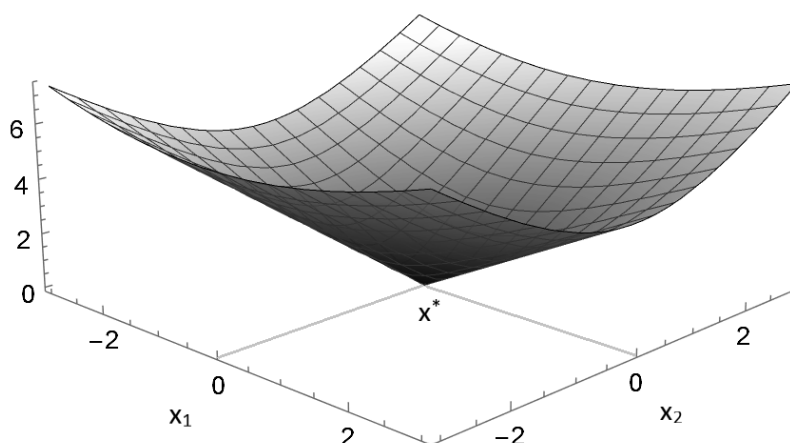
3. Funkciju  $f$  sauc par stingri izliektu uz leju, ja punktā 1) minētā nevienādība izpildās stingrā nozīmē jeb " $<$ ". Toties funkciju  $f$  sauc par stingri izliektu uz augšu, ja punktā 2) minētā nevienādība ir stingra jeb " $>$ " ([1]).

### Piemērs 3.1.3.

Funkcija

$$f(x_1, x_2) = \sqrt{2x_1^2 + 4x_2^2} \quad (3.1.1.)$$

ir izliekta un visur diferencējama, izņemot minimuma punktā  $x^* = (0, 0)^T$  (skatīt 3.1.1. att.).



3.1.1. att. Funkcijas (3.1) grafiks.

Lai atrastu minimuma punktu  $x^*$ , tiek veikta gradienta metode, izvēloties patvaļīgu sākumpunktu  $x^0 = (1; 0, 5)^T$ . Gradients funkcijai  $f(x_1, x_2)$  ir

$$\text{grad } f(x_1, x_2) = \left( \frac{2x_1}{\sqrt{2x_1^2 + 4x_2^2}}; \frac{4x_2}{\sqrt{2x_1^2 + 4x_2^2}} \right)^T.$$

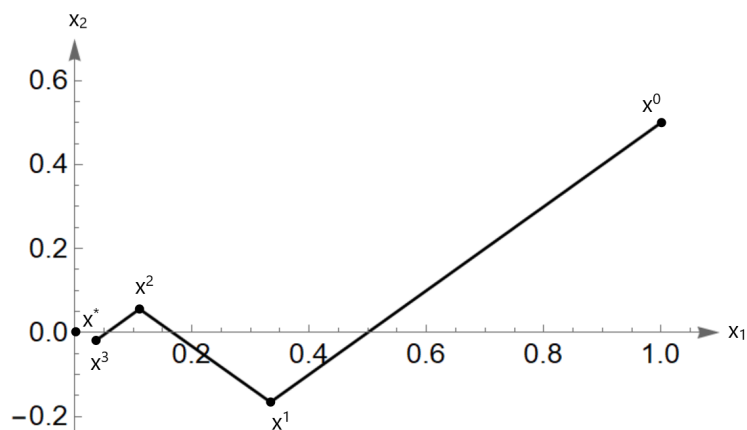
Pirmās iterācijas soļa garumu nosaka, izveidojot palīgfunkciju  $g_0(\alpha)$  un to minimējot, tas ir,

$$\begin{aligned} x^1 &= x^0 - \alpha_0 \text{grad} f(x^0) \approx \\ &\approx (1; 0, 5)^T - \alpha_0 (1, 1547; 1, 1547)^T = \\ &= (1 - 1, 1547\alpha_0; 0, 5 - 1, 1547\alpha_0)^T, \\ g_0(\alpha) &= \sqrt{2(1 - 1, 1547\alpha_0)^2 + 4(0, 5 - 1, 1547\alpha_0)^2} \rightarrow \min \\ g_0'(\alpha) &= \frac{1}{2} * \frac{4(1 - 1, 1547\alpha_0)(-1, 1547) + 8(0, 5 - 1, 1547\alpha_0)(-1, 1547)}{\sqrt{2(1 - 1, 1547\alpha_0)^2 + 4(0, 5 - 1, 1547\alpha_0)^2}} = 0 \Rightarrow \\ &\Rightarrow 2(-1, 1547)(1 - 1, 1547\alpha_0 + 1 - 2, 3094\alpha_0) = 0 \Rightarrow \\ &\Rightarrow \alpha_0 = \frac{2}{3, 4641} \approx 0, 577. \end{aligned}$$

Iegūstot pirmās iterācijas soļa garumu, mēs varam aprēķināt gradienta metodes pirmās iterācijas punktu  $x^1$ .

$$\begin{aligned} x^1 &= (1; 0, 5)^T - 0,577(1, 1547; 1, 1547)^T = \\ &= (0, 334; -0, 166)^T. \end{aligned}$$

Šādi turpinot gradienta metode patvaļīgam sākumpunktam konverģē uz  $x^*$  (skatīt 3.1.2. att.).



3.1.2. att. Funkcijas (3.1) minimuma meklēšana ar gradienta metodi, tuvināto atrisinājumu virkne.

Galvenā problēma ir noteikt atbilstošu apstāšanās kritēriju. Par cik

$$\text{grad}f(x_1, x_2) = \frac{1}{\sqrt{2x_1^2 + 4x_2^2}} \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix},$$

mēs iegūstam, ka  $|\text{grad}f(x_1, x_2)| = \frac{1}{\sqrt{2x_1^2 + 4x_2^2}} \sqrt{4x_1^2 + 16x_2^2} \geq 1$  jebkuram punktam  $(x_1, x_2)^T \neq (0, 0)^T$ , tas ir, neviens no algoritma konstruētajiem punktiem neapmierina apstāšanās kritēriju  $|\text{grad}f(x_1, x_2)| \leq \varepsilon$  ([7]). ■

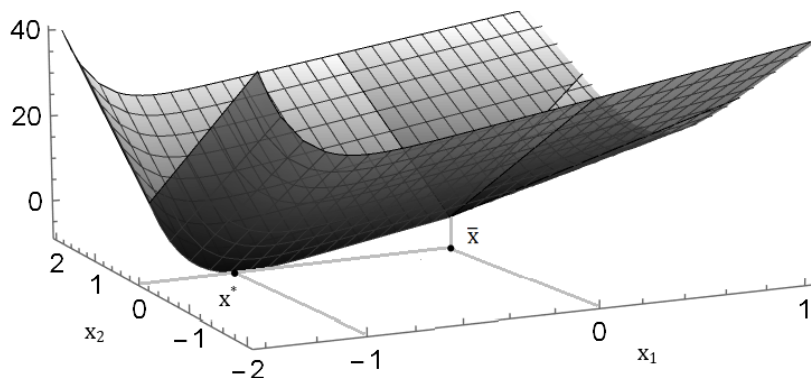
#### Piemērs 3.1.4.

Funkcija

$$f(x_1, x_2) = \begin{cases} 5\sqrt{9x_1^2 + 16x_2^2}, & x_1 \geq |x_2| \\ 9x_1 + 16|x_2|, & 0 < x_1 < |x_2| \\ 9x_1 + 16|x_2| - x_1^9, & x_1 \leq 0 \end{cases} \quad (3.1.2)$$

ir izliekta un diferencējama visos punktos, izņemot staru, kas definēts kā  $x_1 \leq 0, x_2 = 0$  (skatīt 3.1.3. att.). Pirmās funkcijas  $f_1(x_1, x_2) = 5\sqrt{9x_1^2 + 16x_2^2}$ ,  $x_1 \geq |x_2|$ , minimuma punkts ir  $(0;0)$ , otrajai funkcijai  $f_1(x_1, x_2) = 9x_1 + 16|x_2|$ ,  $0 < x_1 < |x_2|$ , minimums neek-

sistē, bet trešās funkcijas  $f_3(x_1, x_2) = 9x_1 + 16|x_2| - x_1^9$  mazākā vērtība tiks sasniegta pie  $x_2 = 0$ ,  $f_3(x_1, 0) = 9x_1 + x_1^9$  minimums būs tad, kad  $(9x_1 - x_1^8)' = 9 - 9x_1^8 = 9(1 - x_1^8) = 0 \Rightarrow x_1^8 = 1 \Rightarrow x_1 = \pm 1$ . Minimums ir pie  $x_1 = 1$ . Dotās funkcijas  $f$  globālais minimums atrodas punktā  $x^* = (-1; 0)^T$ , kas atrodas tieši uz tā stara, kur  $f$  nav diferencējama.



3.1.3. att. Funkcijas (3.2) grafiks.

Veicot minimizāciju ar gradienta metodi, tiek izvēlēts punkts no apgabala

$$R := \{(x_1, x_2)^T \mid x_1 > |x_2| > (\frac{9}{16})^2 x_1\},$$

piemēram,  $x^0 = (1; 0, 5)^T$  (skatīt 3.1.4. att.). Par cik izvēlētajam sākumpunktam  $x^0$  izpildās nevienādība  $x_1 \geq |x_2|$ , tad pirmajā iterācijā gradienta metode tiek rēķināta funkcijai  $5\sqrt{9x_1^2 + 16x_2^2}$ . Gradients funkcijai  $5\sqrt{9x_1^2 + 16x_2^2}$  ir

$$\text{grad} (5\sqrt{9x_1^2 + 16x_2^2}) = \left( \frac{45x_1}{\sqrt{9x_1^2 + 16x_2^2}}; \frac{80x_2}{\sqrt{9x_1^2 + 16x_2^2}} \right)^T.$$

Identiski kā Piemērā 3.1.3., pirmās iterācijas soļa garumu nosaka, izveidojot palīgfunkciju  $g_0(\alpha)$  un to minimizējot, tas ir,

$$\begin{aligned} x^1 &= x^0 - \alpha_0 \text{grad}f(x^0) \approx \\ &\approx (1; 0, 5)^T - \alpha_0(12, 481; 11, 094)^T = \\ &= (1 - 12, 481\alpha_0; 0, 5 - 11, 094\alpha_0)^T, \\ g_0(\alpha) &= 5\sqrt{9(1 - 12, 481\alpha_0)^2 + 16(0, 5 - 11, 094\alpha_0)^2} \rightarrow \min \\ &\alpha_0 = 0, 06. \end{aligned}$$

Pēc tam tiek aprēķināts gradienta metodes pirmās iterācijas punkts  $x^1$ .

$$x^1 = (1; 0, 5)^T - 0, 06(12, 481; 11, 094)^T =$$

$$= (0, 251; -0, 166)^T.$$

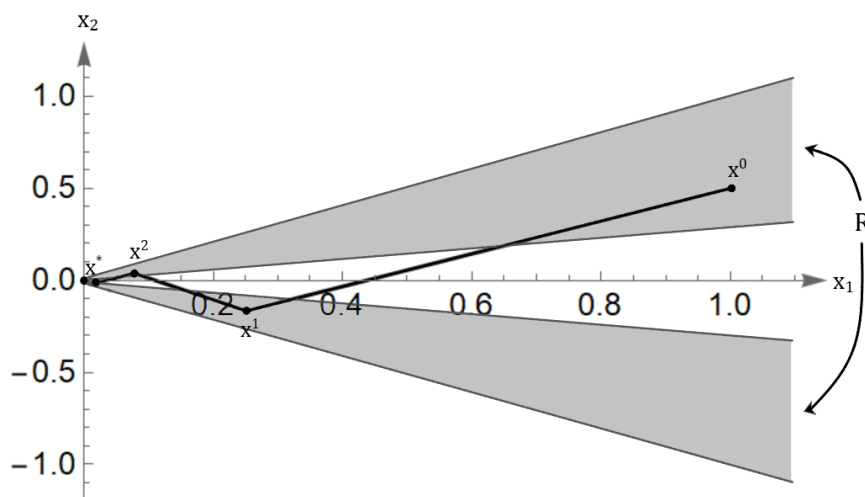
Iterācijas punkts  $x^1$  pieder apgabalam  $R$  un tam arī izpildās nevienādība  $x_1 \geq |x_2|$  līdz ar to  $x^2$  tiek iegūts identiski  $x^1$ , tas ir,

$$\begin{aligned} x^2 &= x^1 - \alpha_1 \operatorname{grad}f(x^1) \approx \\ &\approx (0, 251; -0, 166)^T - \alpha_1(11,251; -13, 228)^T = \\ &= (0, 251 - 11,251\alpha_1; -0, 166 + 13, 228\alpha_1)^T, \\ g_1(\alpha) &= 5\sqrt{9(0, 251 - 11,251\alpha_1)^2 + 16(-0, 166 + 13, 228\alpha_1)^2} \rightarrow \min \end{aligned}$$

$$\alpha_1 = 0, 0154 ,$$

$$\begin{aligned} x^2 &= (0, 251; -0, 166)^T - 0, 0154(11,251; -13, 228)^T = \\ &= (0, 078; 0, 038)^T. \end{aligned}$$

Kā redzam, tad arī otrais iterācijas punkts  $x^2 \in R$  un tam izpildās nevienādība  $x_1 \geq |x_2|$ . Šādi turpinot, iterācijas punktu virkne vienmēr paliks apgabalā  $R$  un konverģēs uz  $\bar{x} := (0, 0)^T$  (skatīt 3.1.5. att.). Tomēr šis punkts nav pat funkcijas  $f$  lokālais minimums (skatīt 3.1.3. att. un 3.1.6. att.).



**3.1.4. att. Funkcijas (3.1.2) tuvināto atrisinājumu virkne un apgabala  $R$  interpretācija.**

Šoreiz izvēlēsimies punktu, kas nepieder apgabalam  $R$ , piemēram,  $x^0 = (1; -0, 1)^T$ . Arī šoreiz izvēlētajam sākumpunktam  $x^0$  izpildās nevienādība  $x_1 \geq |x_2|$ , kas nozīmē, ka atkal pirmajā iterācijā gradienta metode tiek rēķināta funkcijai  $5\sqrt{9x_1^2 + 16x_2^2}$ .

$$x^1 = x^0 - \alpha_0 \operatorname{grad}f(x^0) \approx$$

$$\begin{aligned}
&\approx (1; -0, 1)^T - \alpha_0(14, 868; 2, 643)^T = \\
&= (1 - 14, 868\alpha_0; -0, 1 - 2, 643\alpha_0)^T, \\
g_0(\alpha) &= 5\sqrt{9(1 - 14, 868\alpha_0)^2 + 16(-0, 1 - 2, 643\alpha_0)^2} \rightarrow \min \\
\alpha_0 &= 0, 062, \\
x^1 &= (1; -0, 1)^T - 0, 062(14, 868; 2, 643)^T = \\
&= (0, 078; -0, 264)^T.
\end{aligned}$$

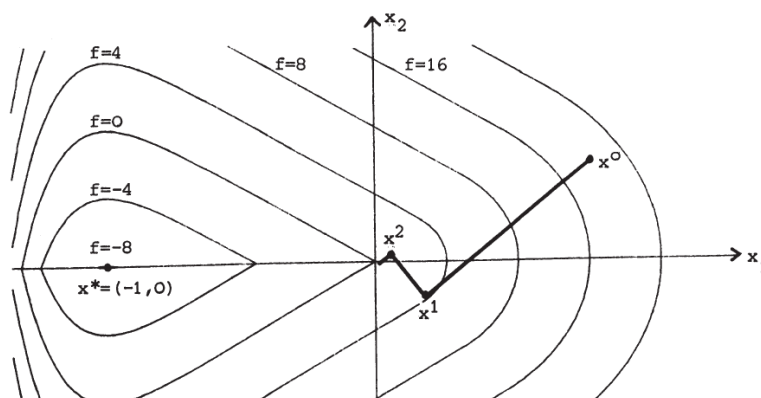
Iterācijas punktam  $x^1$  izpildās nevienādība  $0 < x_1 < |x_2|$  līdz ar to  $x^2$  tiks iegūts, gradienta metodi pielietojot funkcijai  $9x_1 + 16|x_2|$ . Gradients funkcijai  $9x_1 + 16|x_2|$  pie negatīva  $x_2$  ir

$$\text{grad}(9x_1 + 16|x_2|) = (9; -16)^T.$$

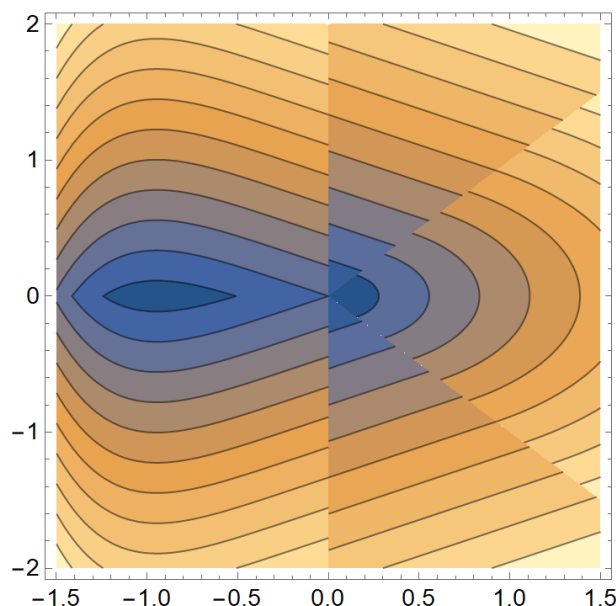
Šoreiz punktu  $x^2$  mēs nevaram iegūt ar identisku pieeju, kā tika iegūts punkts  $x^1$ , jo funkcijai  $g_1(\alpha) = 9(0,078 - 9\alpha_1) - 16(-0,264 + 16\alpha_1)$  nav minimuma. Tas liek meklēt jaunu veidu, kā izvēlēties iterācijas soļa garumu. Izvēloties patvaļīgu soļa garumu, piemēram, 0,0153, mēs varam aprēķināt otro iterācijas punktu  $x^2$ .

$$\begin{aligned}
x^2 &= x^1 - \alpha_1 \text{grad}f(x^1) \approx \\
x^2 &= (0, 078; -0, 264)^T - 0, 0153(9; -16)^T = \\
&= (-0, 0597; -0, 0192)^T.
\end{aligned}$$

Iterācijas punktam  $x^2$  izpildās nevienādība  $x_1 \leq 0$ , kas nozīmē, ka  $x^3$  tiks iegūts, gradienta metodi pielietojot funkcijai  $9x_1 + 16|x_2| - x_1^9$ . Arī šoreiz un turpmākajos iterācijas soļos problēma ir atrast iterācijas soļa garumu. Šādi turpinot gradienta metode lēnām virzās uz globālā minimuma punkta  $x^* = (-1; 0)^T$  pusi ([7,8]). ■



3.1.5. att. Funkcijas (3.1.2) tuvināto atrisinājumu virkne, kad iterācijas sākumpunktu izvēlās apgabalā R.



3.1.6. att. Funkcijas (3.1.2) kontūru grafiks.

## 3.2. Nediferencējamas funkcijas aproksimēšana ar diferencējamu funkciju

Pēdējie divi piemēri parāda, ja funkcija ir pat diferencējama visos iterāciju punktos, gradienta metode var nespēt minimizēt nediferencējamu funkciju.

Vienkārša pieeja, lai minimizētu nediferencējamu funkciju  $f$ , ir balstīta uz  $f$  tuvinājumu ar diferencējamu funkciju  $g$ . Piemēram, funkcija

$$f(x) = |x|$$

var tikt aizstāta ar

$$g(x) = \sqrt{x^2 + \varepsilon},$$

kur  $\varepsilon > 0$  ir "maza" vērtība. Tādā gadījumā, summa

$$f(x) = f_0(x) + \sum_{i=1}^m |f_i(x)| \tag{3.2.3}$$

ar funkcijām  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 0, 1, \dots, m$ , var tikt aproksimēta ar funkciju

$$g(x) = f_0(x) + \sum_{i=1}^m \sqrt{[f_i(x)]^2 + \varepsilon}, \tag{3.2.4}$$

kura ir diferencējama, ja visas funkcijas  $f_i$  ir diferencējamas.

Daudzas nediferencējamas funkcijas var tikt uzrakstītas formā (3.2.3). Piemēram, jeb-

kura nepārtraukta gabaliem lineāra funkcija  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  var tikt izteikta kā

$$f(x) = a + bx + \sum_{i=1}^m \alpha_i |x - x_i| \quad (3.2.5)$$

$$x_1 < x_2 < \dots < x_m; a, b, \alpha_i \in \mathbb{R}, i = 1, \dots, m,$$

kas ir (3.2.3) speciālgadījums pie  $f_0(x) = a + bx$  un  $f_i(x) = \alpha_i(x - x_i)$ ,  $i = 1, \dots, m$ . Tad diferencējama aproksimācija funkcijai (3.2.5) ir formā

$$g(x) = a + bx + \sum_{i=1}^m \sqrt{[\alpha_i(x - x_i)]^2 + \varepsilon} \quad (3.2.6)$$

([7]).

### Piemērs 3.2.1.

Funkcija

$$f(x) = \begin{cases} -\frac{4}{3}x + 5, & x \leq 3 \\ 2 - \frac{1}{3}x, & 3 < x \leq 6 \\ x - 6, & 6 < x \leq 8 \\ 2x - 14, & x > 8 \end{cases} \quad (3.2.7)$$

ir nepārtraukta, izliekta un gabaliem lineāra (skatīt 3.2.1. att.). Vispirms pierakstīsim funkciju formā (3.2.5). Par cik nediferencējami punkti ir 3, 6 un 8, mēs iegūstam

$$f(x) = a + bx + \alpha_1 |x - 3| + \alpha_2 |x - 6| + \alpha_3 |x - 8|. \quad (3.2.8)$$

Parametri  $a, b, \alpha_1, \alpha_2, \alpha_3$  var tikt noteikti, pielīdzinot (3.2.7) un (3.2.8). Ja  $x < 3$ , mēs iegūstam

$$a + bx + \alpha_1(3 - x) + \alpha_2(6 - x) + \alpha_3(8 - x) = -\frac{4}{3}x + 5 \Rightarrow$$

$$a + 3\alpha_1 + 6\alpha_2 + 8\alpha_3 = 5 \quad (i)$$

$$b - \alpha_1 - \alpha_2 - \alpha_3 = -\frac{4}{3}. \quad (ii)$$

Analizējot intervālus  $3 < x \leq 6$ ,  $6 < x \leq 8$ ,  $x > 8$ , mēs iegūstam

$$a + bx + \alpha_1(x - 3) + \alpha_2(6 - x) + \alpha_3(8 - x) = -\frac{1}{3}x + 2 \Rightarrow$$

$$a - 3\alpha_1 + 6\alpha_2 + 8\alpha_3 = 2, \quad (\text{iii})$$

$$b + \alpha_1 - \alpha_2 - \alpha_3 = -\frac{1}{3}. \quad (\text{iv})$$

$$a + bx + \alpha_1(x - 3) + \alpha_2(x - 6) + \alpha_3(8 - x) = x - 6 \Rightarrow$$

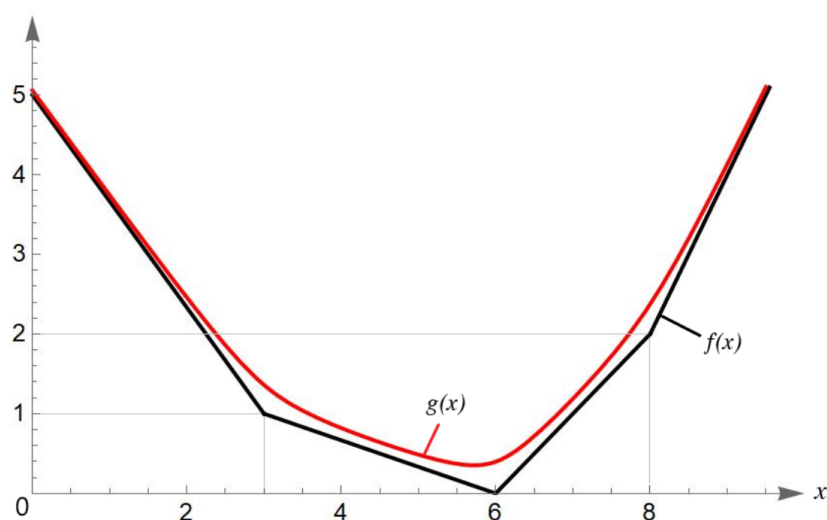
$$a - 3\alpha_1 - 6\alpha_2 + 8\alpha_3 = -6, \quad (\text{v})$$

$$b + \alpha_1 + \alpha_2 - \alpha_3 = 1. \quad (\text{vi})$$

$$a + bx + \alpha_1(x - 3) + \alpha_2(x - 6) + \alpha_3(x - 8) = 2x - 14 \Rightarrow$$

$$a - 3\alpha_1 - 6\alpha_2 - 8\alpha_3 = -14, \quad (\text{vii})$$

$$b + \alpha_1 + \alpha_2 + \alpha_3 = 2. \quad (\text{viii})$$



3.2.1. att. Gabaliem lineāra funkcija  $f$  un tās aproksimācija  $g$ , kas ir diferencējama funkcija ( $\varepsilon = 0.1$ ).

Tā ir lineāru vienādojumu sistēma, kuru var atrisināt. Atiecīgi salīdzinot (ii) un (iv), (iv) un (vi), (vi) un (viii), mēs iegūstam  $\alpha_1 = \frac{1}{2}$ ,  $\alpha_2 = \frac{2}{3}$ ,  $\alpha_3 = \frac{1}{2}$ . Pēc tam vienādojumi (i) un (ii) dod, ka  $a = -\frac{27}{6}$  un  $b = \frac{1}{3}$ . Tādējādi funkcija (3.2.7) var tikt uzrakstīta kā

$$f(x) = \frac{-27}{6} + \frac{1}{3}x + \frac{1}{2}|x - 3| + \frac{2}{3}|x - 6| + \frac{1}{2}|x - 8|, \quad (3.2.9)$$

un saskaņā ar (3.2.6) diferencējamais tuvinājums ir funkcija

$$g(x) = \frac{-27}{6} + \frac{1}{3}x + \sqrt{\left[\frac{1}{2}(x-3)\right]^2 + \varepsilon} + \sqrt{\left[\frac{2}{3}(x-6)\right]^2 + \varepsilon} + \sqrt{\left[\frac{1}{2}(x-8)\right]^2 + \varepsilon} \quad (3.2.10)$$

(skatīt 3.2.1. att.).

3.2.1. tabula

**Funkcijas (3.2.10) minimuma punkta tuvinājumu  $x_\varepsilon$  vērtības pie dažādiem  $\varepsilon$**

| $\varepsilon$ | $x_\varepsilon$ |
|---------------|-----------------|
| 0,1           | 5,72            |
| 0,01          | 5,913           |
| 0,001         | 5,973           |

Kad  $\varepsilon \rightarrow 0$ , punktu virkne  $x_\varepsilon$  konverģē uz funkcijas (3.2.7) minimuma punktu  $x^* = 6$  ([7]). ■

Praksē iepriekš minētā piemērā ieteiktā pieeja rada nopietnas problēmas attiecībā uz procedūras stabilitāti. Tas ir, funkcijas (3.2.3) minimuma punkts parasti apmierina sakarību  $f_k(x^*) = 0$ ,  $k \in \{1, \dots, m\}$  (skatīt 3.3.1. att.). Pie maza  $\varepsilon$  un punktam  $x$ , kas ir tuvu  $x^*$ ,  $k$ -tais summas (3.2.4) loceklis ir maza skaitļa  $z = [f_k(x)]^2 + \varepsilon$  kvadrātsakne. Par cik  $z \rightarrow 0$ , tad funkcijas  $\sqrt{z}$  virziena koeficients konverģē uz  $\infty$ . Tātad sīkas kļūdas  $z$  var radīt lielas kļūdas  $\sqrt{z}$ .

Visbeidzot mēs varam atzīmēt, ka *minimax* uzdevums (2.1) var tikt pārveidots ekvivalentā ierobežojumu formā

$$\min_{f_i(x) \leq x_{n+1}} x_{n+1}, \quad i = 1, \dots, m, \quad (3.2.11)$$

ieviešot vēl vienu citu mainīgo  $x_{n+1}$ . Tātad nediferencējamība tiek izslēgta, ja visas funkcijas  $f_i$  ir diferencējamas. Tomēr šī transformācija "nevienkāršo" uzdevumu.

Vairāki nediferencējami uzdevumu algoritmi cenšas pielāgoties diferencējamu funkciju pamatidejām. Nozīmīga metode ir saišķa ("bundle") metode, plakņu griešanas un elipsoīdu metodes ([7]).

### 3.3. Saišķa (*bundle*) metodes pamatideja

Lai gan lielākā daļa šīs teorijas paliek spēkā zem vājākiem nosacījumiem, tālāk mēs pētīsim funkcijas  $f$  minimizēšanu, pie nosacījuma, ka  $f$  ir izliekta telpā  $\mathbb{R}^n$ . Izliektība

nodrošina, ka vienpusējs atvasinājums  $D^+f(x, d)$  ir definēts katrā punktā  $x$  un katram virzienam  $d$ .

Pieņemsim, ka dota funkcija  $f : M \rightarrow \mathbb{R}$ ,  $M \subset \mathbb{R}^n$  un  $x^* \in M$  ir kopas  $M$  iekšējs punkts. Jebkuram vektoram  $y \in \mathbb{R} \setminus \{0\}$  tiek apskatīta viena argumenta funkcija

$$f_{(y)}(t) = f(x^T + ty).$$

Funkcijas  $f$  atvasinājums noteiktā virzienā var tikt interpretēts kā šīs funkcijas atvasinājums pēc  $t$  punktā  $t = 0$ .

**Definīcija 3.3.1.** ([7]) Par funkcijas  $f$  vienpusējo atvasinājumu  $Df(x^*, y)$  virzienā  $y$  sauc robežu

$$Df(x^*, y) = f'_{(y)} = \lim_{t \rightarrow 0} \frac{f_{(y)}(t) - f_{(y)}(0)}{t}.$$

Par funkcijas  $f$  vienpusējo labās puses atvasinājumu  $D^+f(x^*, y)$  un kreisās puses atvasinājumu  $D^-f(x^*, y)$  virzienā  $y$  sauc atbilstošo robežu

$$D^+f(x^*, y) = \lim_{t \rightarrow 0^+} \frac{f_{(y)}(t) - f_{(y)}(0)}{t},$$

$$D^-f(x^*, y) = \lim_{t \rightarrow 0^-} \frac{f_{(y)}(t) - f_{(y)}(0)}{t}.$$

**Piemērs.** Pieņemsim, ka dota funkcija  $f(x_1, x_2) = x_1^2 + x_1x_2$ ,  $M = \mathbb{R}^2$ . Pieņemsim, ka  $x^* = (1; 1)$  un  $y = (2; 3)$ , tad

$$f_{(y)}(t) = f(x^* + ty) = f(1 + 2t; 1 + 3t) = (1 + 2t)^2 + (1 + 2t) * (1 + 3t).$$

Tā kā

$$f'_{(y)}(t) = 2 * (1 + 2t) * 2 + 2 * (1 + 3t) + (1 + 2t) * 3,$$

tad  $Df(x^*, y) = f'_{(y)}(0) = 2 * 1 * 2 + 2 * 1 + 1 * 3 = 4 + 2 + 3 = 9$ . Vienpusējie atvasinājumi šajā gadījumā sakrīt ar 9. ■

**Teorēma 3.3.2.** ([7]) Ja tiek dota izliekta funkcija  $f : M \rightarrow \mathbb{R}$  ( $M \subset \mathbb{R}$  ir izliekta) un iekšējais punkts  $x^* \in M$ , tad tai izpildās

$$-D^+f(x^*; -y) = D^-f(x^*, y) \quad \forall y \in \mathbb{R}^n. \quad (i)$$

$$D^+f(x^*; \lambda y) = \lambda D^+f(x^*, y) \text{ un } D^-f(x^*; \lambda y) = \lambda D^-f(x^*, y), \text{ kad } \lambda \geq 0 \quad \forall y \in \mathbb{R}^n. \quad (ii)$$

Vektors  $a \in \mathbb{R}^n$  ir  $f$  apakšgradients punktā  $x^*$  tad un tikai tad, ja

$$D^+ f(x^*; y) \geq a^T y \quad \forall y \in \mathbb{R}^n. \quad (\text{iii})$$

Pirmā ideja ir minimizēt  $f$  iteratīvi, modificējot gradienta metodi nediferencējamai funkcijai  $f$ . Šajā gadījumā ātrākais virziena koeficients  $d^k$  tiek iegūts, atrisinot uzdevumu

$$\min_{\|d\|=1} D^+ f(x^k, d), \quad (3.3.1)$$

jo priekš  $\|d\| = 1$  atvasinājums  $D^+ f(x^k, d)$  ir  $f$  pieauguma vai samazinājuma mērs virzienā  $d$ . Uzdevums (3.3.1) ir ekvivalents ar uzdevumu

$$\min_{\|d\| \leq 1} D^+ f(x^k, d), \quad (3.3.2)$$

ciktāl optimālie atrisinājumi ir identiski. Tomēr pēdējais skaitļošanas ziņā ir vieglāks, jo pieļaujamā kopa ir izliekta. Saskaņā ar Teorēmu 3.3.2 uzdevumu (3.3.2) var pārveidot formā

$$D^+ f(x, d) = \max_{a \in \partial f(x)} a^T d, \quad (3.3.3)$$

kur  $\partial f(x)$  apzīmē  $f$  apakšdiferenciāli punktā  $x$  (Definīcija 1.4.5.). Savienojot (3.3.2) un (3.3.3), mēs iegūstam uzdevumu

$$\min_{\|d\| \leq 1} \max_{a \in \partial f(x)} a^T d, \quad (3.3.4)$$

kuru var pārraksīt ar von Neimana "minimax" teorēmas palīdzību kā

$$\max_{a \in \partial f(x)} \min_{\|d\| \leq 1} a^T d. \quad (3.3.5)$$

Tā kā  $\frac{-a}{\|a\|}$  ir "iekšējā" uzdevuma atrisinājums, mēs iegūstam

$$\max_{a \in \partial f(x)} a^T \frac{-a}{\|a\|} = \max_{a \in \partial f(x)} -\|a\|, \quad (3.3.6)$$

kas beidzot ir ekvivalents "minimum-normas" uzdevumam

$$\min_{a \in \partial f(x)} \|a\|. \quad (3.3.7)$$

Izmantojot izteiksmi (3.3.7), lai noteiktu meklēšanas virzienu, mēs iegūstam sekojošu pamatprocedūru:

- (1) Izvēlās  $x_0 \in \mathbb{R}^n$  un nosaka, ka  $k = 0$ .
- (2) Izrēķina  $d^k = \frac{-a^k}{\|a^k\|}$ , kur  $a^k$  ir uzdevuma  $\min_{a \in \partial f(x^k)} \|a\|$  atrisinājums.
- (3) Ja  $0 \in \partial f(x^k)$ , tad ir jāapstājas. Citādi ir jānodod uz procedūras soli (4).
- (4) Jānosaka  $\min_{\lambda \in \mathbb{R}_+} f(x^k + \lambda d^k)$  atrisinājums  $\lambda_k$ .
- (5) Nosaka  $x^{k+1} = x^k + \lambda_k d^k$ ,  $k = k + 1$  un dodās uz procedūras soli (2).

Apstāšanās kritērijs solim (3) ir balstīts uz faktu, ka  $x^k$  ir  $f$  globālais minimuma punkts tad un tikai tad, ja  $0 \in \partial f(x^k)$ . Diferencējamajos punktos iterācija ir identiska gradientu metodes solim ([7]).

### Piemērs 3.3.3.

Mēs minimizēsim funkciju

$$f(x_1, x_2) = |x_1| + |x_2|,$$

sākumpunktu izvēlamies kā  $x^0 = (3; -2)^T$ .  $f$  apakšgradients  $a = (a_1; a_2)^T$  nediferencējamā punktā  $x^0$  apmierina nevienādību:

$$|x_1| + |x_2| \geq 5 + (a_1; a_2) \begin{pmatrix} x_1 - 3 \\ x_2 + 2 \end{pmatrix}.$$

No tā seko vienādojumu sistēma:

$$f(x) = \begin{cases} |x_1| \geq a_1 x_1, & \Rightarrow |a_1| \leq 1 \\ |x_2| \geq a_2 x_2, & \Rightarrow |a_2| \leq 1 \\ 0 \geq 5 - 3a_1 + 2a_2, & \Rightarrow a_1 \geq \frac{5+2a_2}{3}. \end{cases}$$

Atrisinot vienādojumu sistēmu mēs iegūstam, ka vektors  $a$  ir apakšgradients tad un tikai tad, ja  $a_1 = 1$  un  $a_2 = -1$ , tas ir,  $f$  apakšdiferenciālis  $x^0$  ir formā

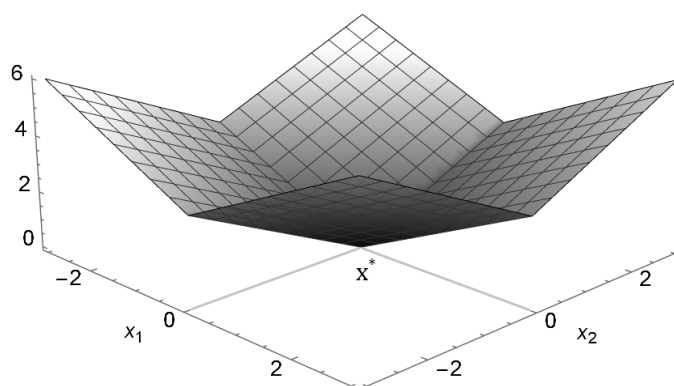
$$\partial f(x^0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Algoritma otrais solis (2) dod  $a^0 = (1; -1)^T$  un  $d^0 = \frac{-(1; -1)^T}{\|(1; -1)\|} = \frac{-(1; -1)^T}{\sqrt{2}} = (-\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}})^T$ . Algoritma (4) solī mēs nosakām soļa garumu  $\lambda$ .

$$\begin{aligned} \min_{\lambda \in \mathbb{R}_+} f(x^0 + \lambda d^0) &= \min_{\lambda \in \mathbb{R}_+} f(3 - \frac{1}{\sqrt{2}}\lambda; -2 + \frac{1}{\sqrt{2}}\lambda)^T = \\ &= \min_{\lambda \in \mathbb{R}_+} |3 - \frac{1}{\sqrt{2}}\lambda| + |-2 + \frac{1}{\sqrt{2}}\lambda|. \end{aligned}$$

Mazākā vērtība tiek sasniegta intervālā, kad  $\lambda \in [2, 83; 4, 243]$ . No iepriekš minētā intervāla patvaļīgi izvēlās, ka  $\lambda_0 = 3$ . Turpinājumā veicot soli (5), tiek atrasts punkts  $x^1$ , tas ir,

$$x^1 = x^0 + \lambda_0 d^0 = (3; -2)^T + 3 * (-\frac{1}{\sqrt{2}}; \frac{1}{\sqrt{2}})^T \approx (0, 879; 0, 121)^T.$$



3.3.1. att. Piemēra 3.3.1. funkcijas grafiks.

Dodamies uz soli (2) un nosakām apakšgradientu punktā  $(0, 879; 0, 121)^T$ .  $f$  apakšgradients  $a = (a_1; a_2)^T$  nediferencējamā punktā  $x^1$  apmierina nevienādību:

$$|x_1| + |x_2| \geq 1 + (a_1; a_2) \begin{pmatrix} x_1 - 0, 879 \\ x_2 - 0, 121 \end{pmatrix}.$$

No tā seko vienādojumu sistēma:

$$f(x) = \begin{cases} |x_1| \geq a_1 x_1, & \Rightarrow |a_1| \leq 1 \\ |x_2| \geq a_2 x_2, & \Rightarrow |a_2| \leq 1 \\ 0 \geq 1 - 0, 879 a_1 - 0, 121 a_2, & \Rightarrow a_1 \geq \frac{1 - 0, 121 a_2}{0, 879}. \end{cases}$$

No vienādojumu sistēmas tiek noteikts, ka vektors  $a$  ir apakšgradients tad un tikai tad, ja  $a_1 = 1$  un  $0, 064 \leq a_2 \leq 1$ , tas ir,  $f$  apakšdiferenciālis punktā  $x^1$  ir

$$\partial f(x^1) = \left\{ \begin{pmatrix} 1 \\ \alpha \end{pmatrix} \mid 0, 064 \leq \alpha \leq 1 \right\}.$$

Algoritma otrais solis (2) dod  $a^1 = (1; 0, 064)^T$  un  $d^1 = (-0, 998; -0, 064)^T$ . Pēc tam turpinam ar algoritma (4) soli.

$$\begin{aligned} \min_{\lambda \in \mathbb{R}_+} f(x^1 + \lambda_1 d^1) &= \min_{\lambda \in \mathbb{R}_+} f(0, 879 - 0, 998\lambda; 0, 121 - 0, 064\lambda)^T = \\ &= \min_{\lambda \in \mathbb{R}_+} |0, 879 - 0, 998\lambda| + |0, 121 - 0, 064\lambda|. \end{aligned}$$

Mazākā vērtība tiks sasniegta ar  $\lambda_1 = 1$ . Pēc tam algoritma (5) solī atrodam punktu  $x^2$

$$x^2 = x^1 + \lambda_1 d^1 = (0, 879; 0, 121)^T + 0,881 * (-0,998; -0,064)^T \approx (0; 0,065)^T.$$

Šādi turpinot jau nākamajā solī jeb trešajā iterācijā varam teikt, ka algoritms konverģē uz funkcijas minimumu un punktā  $x^3$  tiek apmierināts algoritma apstāšanās kritērijs  $0 \in \partial f(x^3)$ , kas liecina, ka ir atrasts funkcijas  $f$  minimums  $(0; 0)^T$  (skatīt 3.3.1 att.) ([7]). ■

#### Piemērs 3.3.4.

Izvēloties sākumpunktu  $x^0 = (0, 0)^T$ , mēs izmantosim algoritmu funkcijas (3.1.2) minimizācijai. Vienkāršības dēļ mēs noteiksim ātrāko virzienu tieši ar (3.3.1) palīdzību. Tā kā

$$D^+ f(x^0, d) = \lim_{t \rightarrow 0^+} \frac{f(td)}{t},$$

$$f(td) = \begin{cases} 5t\sqrt{9d_1^2 + 16d_2^2}, & d_1 \geq |d_2|, \\ t(9d_1 + 16|d_2|), & 0 < d_1 < |d_2|, \\ 9td_1 + 16t|d_2| - (td_1)^9, & d_1 \leq 0, \end{cases}$$

mēs iegūstam

$$D^+ f(x^0, d) = \begin{cases} 5\sqrt{9d_1^2 + 16d_2^2}, & d_1 \geq |d_2|, \\ 9d_1 + 16|d_2|, & d_1 < |d_2|, \end{cases}$$

un (3.3.1) atrisinājums ir  $d^0 = (-1, 0)^T$  ar optimālo vērtību -9. Attēls 3.1.3. ilustrē to, ka  $d^0$  patiešām ir ātrākais virziens. Algoritma (4) solī tiek iegūts pirmās iterācijas soļa garums  $\lambda_0$ , t.i.,

$$\begin{aligned} \min_{\lambda \in \mathbb{R}_+} f(x^0 + \lambda_0 d^0) &= \min_{\lambda \in \mathbb{R}_+} f(-\lambda; 0)^T = \\ &= \min_{\lambda \in \mathbb{R}_+} \lambda^9 - 9\lambda. \end{aligned}$$

Mazākā vērtība tiks sasniegta ar  $\lambda_0 = 1$ . Tālāk algoritma (5) solī iegūstam punktu  $x^1$

$$x^1 = x^0 + \lambda_0 d^0 = (0; 0)^T + 1 * (-1; -0)^T \approx (-1; 0)^T.$$

Par cik algoritma apstāšanās kritērijs  $0 \in \partial f(x^1)$  ir apmierināts, tad varam teikt, ka optimālais atrisinājums ir atrasts pirmajā iterācijā ([7]). ■

Iepriekš minētais algoritms ne vienmēr konverģē uz optimālo atrisinājumu. Piemēram, ja tas ir pielietots funkcijai (3.1.2) ar sākuma punktu piemēra 3.1.4. reģionā  $R$ , tad virkne  $x^k$  konverģē uz neoptimālo atrisinājumu  $\bar{x} = (0, 0)^T$  (skatīt 3.1.3. att. un 3.1.5. att.). Par cik problēma (3.1.2) ir diferencējama pie  $x_1 > 0$ , iepriekš minētais algoritms ģenerē tos pašu iterāciju punktus kā gradienta metode piemērā 3.1.2. Tomēr piemērs 3.1.2. ierosina pamatalgoritma sekojošo modifikāciju.

Punktam  $x^k$  atbilstošo apkārtni  $U_k$  mēs solī (2) aizstājam apakšdiferenciāli  $\partial f(x^k)$  ar apakšdiferenciāļu apvienojumu  $D(U_k) := \bigcup_{y \in U_k} \partial f(y)$ . Ja punkts  $x^k$  ir pietiekami tuvu  $\bar{x}$ , tad  $\bar{x} \in U_k$  un  $\partial f(\bar{x} \subset D(U_k))$ , tas ir, ideālais virziens ir  $(-1, 0)^T$ , kas atbilst punk-

tam  $\bar{x}$  un kļūst par virziena meklēšanas kandidātu. Izvēloties šo meklēšanas virzienu, modificētais algoritms var "izbēgt no pievilkšanas punkta"  $\bar{x}$ .

Komerciālajās programmatūrās "bundle" metodē virziena meklēšana ir nosakāma sarežģītā veidā. Šeit mēs paturpināsim ar dažām pamatidejām. Apakšgradienti  $g_1, g_2, g_3, \dots$ , kas ir noteikti iepriekšējās iterācijās tā, ka labs tuvinājums  $D(U_k)$  ir izliekts daudzskaldnis  $P := \text{conv}(\{g_1, g_2, g_3, \dots\})$ .  $P$  apakš- gradients norāda tā saukto "bundle". Meklēšanas virziena noteikšana ar  $P$  palīdzību ir kvadrātiskās minimizācijas uzdevums, kuru var viegli atrisināt. Ja iegūtais virziens ir "labs" (līnijas meklēšana to norādīs), tad algoritms virzās šajā virzienā. Pretējā gadījumā  $D(U_k)$  labāks tuvinājums tiek veidots, iekļaujot citu apakšgradientu daudzskaldnī  $P$  ([7]).

## Secinājumi

Pēc bakalaura darba izpildes un izpētītajām nediferencējamu funkciju optimizācijas metodēm var secināt, ka

1. Nediferencējamu funkciju optimizācija ir tikai savā attīstības sākuma posmā, kur nezināmo ir vairāk par zināmo. Vēl nav izveidota neviena optimizēšanas metode, kas pilnībā derētu katrai nediferencējamai funkcijai. Tas ļoti sarežģī darbu dažādu nozaru speciālistiem, kuriem būtu nepieciešama ātra un efektīva metode, lai ietaupot laiku, veiktu sev nepieciešamos optimizācijas uzdevumus.
2. Gradients metode ir efektīvs veids, kā minimizēt nediferencējamu funkciju līdz brīdim, kad funkcijai ir vairāki lokālie minimuma punkti vai ir nepieciešams nospraust iterāciju apstāšanās sliekšni. Ne vienmēr pieminētais apstāšanās sliekšnis darbojas. Vairāku lokālo minimumu gadījumā iterācijas sākumpunkta izvēle spēlē ļoti nozīmīgu lomu, jo tā nepareiza izvēle ietekmēs vai iterācijas punkti tuvosies lokālajam minimumam vai globālajam minimumam. Nedrīkst arī piemirst iterācijas soļa izvēles nozīmīgumu, jo tas ietekmēs, cik ātri gradienta metode konverģēs uz ekstrēma punktu.
3. Nediferencējamu funkciju aproksimēšana ar diferencējamu funkciju ir vienkāršs veids, kā minimizēt funkciju, aizstājot nediferencējamu funkciju ar diferencējamu. Tas ir efektīvs veids, kā optimizēt izliektu, gabaliem lineāru funkciju. Tomēr šai metodei praksē var rasties nopietnas problēmas ar procedūras stabilitāti, jo tuvojoties atrisinājumam ir jāreķina kvadrātsakne no ļoti maza skaitļa.
4. Saišķa metodes algoritms ir efektīva metode, lai nonāktu pie meklētā minimuma punkta. Šī metode, lai gan arī daļēji balstās uz gradienta metodi, ir daudz ātrāka, jo tiek noteikts precīzāks minimuma punkta virziens. Arī šajai metodei ir trūkumi, tas ir, iespējami lieli sarežģījumi ar apakšdiferenciāļa noteikšanu un problēmas ar iterācijas sākumpunkta izvēli.

Darba turpinājumā noteikti noderētu aplūkot un izpētīt vēl citas nediferencējamu funkciju optimizācijas metodes, kuras šajā darbā netika aplūkotas kā, piemēram, mašīnmācīšanās algoritmus, kuriem ir liels potenciāls, lai izveidotu vienu vienotu optimizācijas metodi visām nediferencējamām funkcijām.

## Izmantotā literatūra un avoti

- [1] I. Bula. *Optimizācijas metodes, lekciju konspekts*. LU e-studijas, 2017.
- [2] C. A. Floudas, P. M. Pardalos. *Encyclopedia of Optimization, 2nd Edition*. Springer, 2009.
- [3] D. Kļaviņš. *Optimizācijas metodes ekonomikā 1., 2.. Otrais izdevums*. Datorzinību Centrs, 2003.
- [4] *Mathematical optimization*. Pieejams [www.en.wikipedia.org](http://www.en.wikipedia.org): [https://en.wikipedia.org/wiki/Mathematical\\_optimization](https://en.wikipedia.org/wiki/Mathematical_optimization), skatīts 16.05.2021.
- [5] N. Robinson. *Nondifferentiable Optimization*. Pieejams: [www.optimization.mccormick.northwestern.edu](http://www.optimization.mccormick.northwestern.edu): [https://optimization.mccormick.northwestern.edu/index.php/Nondifferentiable\\_Optimization](https://optimization.mccormick.northwestern.edu/index.php/Nondifferentiable_Optimization), skatīts 15.05.2021.
- [6] N. A. Sofi, S. Maqbool, B. Bilal, A. Ahmad. *Mathematical Programming: A Brief Historical Sketch and its Applications*. Department of Statistics, University of Kashmir, 2014.
- [7] P. Zoring. *Nonlinear Programing. An Introduction*. De Gruyter, 2014.
- [8] J. Zowe. *Nondifferentiable Optimization*. Raksts no grāmatas: Schittkowski K. *Computational Mathematical Programming*. Springer-Verlag Berlin, Heidelberg, 1985.

Bakalaura darbs „Nediferencējamu funkciju optimizācijas problēmas”  
izstrādāts LU Fizikas, matemātikas un optometrijas fakultātē.

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā  
norādītie informācijas avoti un iesniegtā noslēguma darba elektroniskā versija atbilst  
LUIS augšupielādētā darba elektroniskai kopijai.

Autors: Kristaps Feikners

Rekomendēju darbu aizstāvēšanai

Vadītāja: profesore Dr. mat. Inese Bula

Recenzents: docents Jānis Bajārs

Darbs iesniegts \_\_\_\_\_ nodaļā 03.06.2021.

Dekāna pilnvarotā persona: metodiķe Inita Šneidere

Darbs aizstāvēts bakalaura gala pārbaudījuma komisijas sēdē

\_\_\_\_\_. prot. Nr. \_\_\_\_.

Komisijas sekretārs: lektors Artis Luguzis/asoc. prof. Ingrīda Uljane