Nužni i dovoljni uvjeti optimalnosti

Cestarić, Lucija

SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO–MATEMATIČKI FAKULTET

MATEMATIČKI ODSJEK

Lucija Cestarić

NUŽNI I DOVOLJNI UVJETI OPTIMALNOSTI

Diplomski rad

Voditelj rada:
izv. prof. dr. sc. Marko Vrdoljak

Zagreb, rujan 2019.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ____________________________, predsjednik
2. ____________________________, član
3. ____________________________, član

Povjerenstvo je rad ocijenilo ocjenom _______________.

Potpisi članova povjerenstva:

1. ____________________________
2. ____________________________
3. ____________________________

Sadržaj

Uvod

Optimizacija ima značajnu ulogu u svijetu u kojem živimo. Mnoge stvari koje radimo pokušavamo napraviti na optimalan način, to jest želimo maksimizirati korisnost ili minimizirati trošak. Primjerice, poslovođa želi rasporediti posao između svojih zaposlenika tako da bude obavljen u najkraćem mogućem roku, poljoprivrednik želi odrediti količinu kukuruza i žita koju treba posijati kako bi ostvario maksimalnu dobit, zrakoplovna industrija želi rasporediti rute pojedinih zrakoplova kako bi let trajao što kraće... Matematički gledano, optimizacijski problemi (problemi matematičkog programiranja) sastoje se u određivanju ekstrema (minimuma ili maksimuma) neke funkcije uz zadana ograničenja. To znači da se između brojnih mogućih rješenja određenog problema nastoji pronaći najbolje, to jest optimalno po nekom kriteriju.

Glavni cilj ovog rada je konstruirati kriterije koji nisu teški za provjeriti, a omogućuju nam da ispitamo da li je neka točka koja pripada dopustivom skupu točka lokalnog, odnosno globalnog ekstrema. Jasno je da nužni uvjeti moraju biti zadovoljeni u svakoj točki lokalnog ekstrema, ali mogu ih zadovoljavati i mnoge druge točke koje nemaju veze s optimalnim vrijednostima. Dakle, ispitivanjem nužnih uvjeta možemo samo izdvojiti točke koje su kandidati za točke lokalnih ekstrema. S druge strane, dovoljni uvjeti predstavljaju snažne kriterije koji impliciraju da se radi o točki lokalnog ekstrema.

Rad je podijeljen na tri poglavlja. U prvom poglavlju obrađeno je pitanje egzistencije optimalnog rješenja zadaće nelinearnog programiranja. Uvodni dio poglavlja donosi osnovne definicije i rezultate koji su ključni za razumijevanje materije rada. Uvodimo osnovne pojmove kao što su točke globalnog ili lokalnog minimuma te konveksni skup i konveksna funkcija. Pojam konveksnosti nam je izuzetno važan i već na početku dokazujemo da su u slučaju konveksnih funkcija točke lokalnih minimuma ujedno i točke globalnih minimuma. Osvrnuli smo se na koercitivne funkcije i kao najvažniji rezultat poglavlja navodimo Weierstrassov teorem. Drugo poglavlje je najopsežnije i bavi se razvojem nužnih i dovoljnih uvjeta optimalnosti. Uvodimo različite konuse pridružene dopustivom skupu, koji su ilustrirani na primjerima. Također, dajemo definiciju stacionarne točke u terminu novouvedenog konusa te iskazujemo geometrijske uvjete optimalnosti. Dodatnim uvjetima tipa nejednakosti na dopustivi skup dolazimo do Fritz Johnovih uvjeta optimalnosti, koji su ipak dosta slabi. Zbog toga nastavljamo dalje i pod pretpostavkom da dopustivi skup zadovoljava određene uvjete regularnosti dolazimo do glavnog teorema cijelog rada koji opisuje Karush-Kuhn-Tuckerove uvjete optimalnosti. Oni igraju vrlo važnu ulogu u problemima uvjetne optimizacije. Ukratko, ovi uvjeti govore da dopustiva točka x može biti točka lokalnog minimuma ukoliko ne postoji dopustivi smjer silaska u x. Budući da su uvjeti regularnosti ključni za formiranje KKT uvjeta, posvećujemo im posebnu pozornost. Navodimo neke najpoznatije, od kojih je najčešće korišten uvjet linearne nezavisnosti jer je za probleme koji ga zadovoljavaju rješenje KKT sustava nužno jedinstveno. Na samom kraju poglavlja dokazujemo da su Karush-Kuhn-Tuckerovi uvjeti ne samo nužni, već i dovoljni za lokalnu optimalnost u zadaćama konveksnog programiranja. U posljednjem, trećem poglavlju se bavimo primjenom Karush-Kuhn-Tuckerovih uvjeta na probleme iz područja financija i ekonomije. Kao primjer financijskog problema istaknut je problem izbora portfelja investicija, dok su ekonomske primjene ilustrirane na primjerima određivanja optimalne cijene usluge te maksimizacije prihoda uz ograničenje dobiti.

Poglavlje 1

Egzistencija rješenja zadaće nelinearnog programiranja

Poglavlje započinjemo osnovnom podjelom zadaća matematičkog programiranja te objašnjavanjem pojmova dopustivog skupa i dopustive točke. Zatim ćemo dati formalne definicije globalnog, odnosno lokalnog minimuma te se prisjetiti pojma konveksnog skupa te konveksne funkcije. U drugom dijelu poglavlja definiramo koercitivnu funkciju te dolazimo do osnovnog teorema o egzisteniji optimalne točke.

1.1 Osnovni pojmovi

Opća formulacija zadaće matematičkog programiranja je

$$
\begin{aligned}
\min\ & f(x) \\
& g_i(x) \leq 0, \quad i = 1, \ldots, m, \\
& h_j(x) = 0, \quad j = 1, \ldots, k, \\
& x \in \mathbb{R}^n,
\end{aligned}
$$

pri čemu su $f\colon \mathbb{R}^n \to \mathbb{R}$, $g_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $h_j\colon \mathbb{R}^n \to \mathbb{R}$, $j = 1, \ldots, k$ zadane funkcije te m i $k \in \mathbb{N}_0$. Često se u zadaćama javljaju uvjeti tipa jednakosti i nejednakosti pa govorimo o problemima uvjetne optimizacije. Funkciju f koju želimo minimizirati (ili maksimizirati) nazivamo funkcijom cilja, a funkcije g_i, $i = 1, \ldots, m$ te h_j, $j = 1, \ldots, k$ funkcijama ograničenja. Skup $S = \{ x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \ldots, m;\ h_j(x) = 0, j = 1, \ldots, k \}$ zovemo dopustivi skup, dok za element skupa S kažemo da je dopustiva točka. Ako je funkcija cilja linearna i funkcije ograničenja afine[1] govorimo o zadaći linearnog programi-

[1] Afina funkcija općenito je funkcija $\mathbb{R}^n \to \mathbb{R}$ zadana kao $f(x) = ax + b$. Linearna funkcija je poseban oblik afine za koju je $a \neq 0$ i $b = 0$, to jest funkcija oblika $f(x) = ax$.

ranja. S druge strane, ako je funkcija cilja ili neka od funkcija ograničenja nelinearnog tipa radi se o zadaći nelinearnog programiranja.

Primijetimo da se dovoljno ograničiti na problem minimizacije neke funkcije $f\colon S \subseteq \mathbb{R}^n \to \mathbb{R}$ jer vrijedi

$$\max_{x\in S} f(x) = -\min_{x\in S}(-f(x)).$$

Definicija 1.1. *Neka je $x^* \in S \subseteq \mathbb{R}^n$. Kažemo da je x^** ***točka globalnog minimuma*** *funkcije f na skupu S ako vrijedi*

$$f(x^*) \leq f(x), \quad x \in S.$$

Neka je $K(x^*, r) := \{\, y \in \mathbb{R}^n \mid \|y - x^2\|_2 < r \,\} = \left\{ y \in \mathbb{R}^n \mid \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} < r \right\}$ otvorena kugla oko x^* radijusa r.

Definicija 1.2. *Neka je $x^* \in S \subseteq \mathbb{R}^n$.*

(i) Kažemo da je x^* ***točka lokalnog minimuma*** *funkcije f na skupu S ako vrijedi*

$$\exists r > 0, \forall x \in S \cap K(x^*, r), \quad f(x^*) \leq f(x). \tag{1}$$

(ii) Kažemo da je x^* ***točka strogog lokalnog minimuma*** *funkcije f na skupu S ako vrijedi*

$$\exists r > 0, \forall x \in S \cap K(x^*, r), x \neq x^*, \quad f(x^*) < f(x).$$

Budući da uvjet (1) kaže da je x^* točka lokalnog minimuma ako postoji radijus $r > 0$ takav je x^* globalno optimalna točka na presjeku skupa S i otvorene kugle $K(x^*, r)$, očito je da je globalni minimum ujedno i lokalni. Pitamo se kada vrijedi i obratno, kada je točka za koju funkcija postiže lokalni minimum ujedno i točka za koju se postiže globalni minimum? Odgovor na to nam daje sljedeći teorem, jedan od najvažnijih u optimizaciji. On ističe posebnost konveksnog slučaja pa ćemo se zato prije nego ga iskažemo podsjetiti pojmova konveksnog skupa i konveksne funkcije.

Napomena 1.3. *(i) Kažemo da je skup $S \subseteq \mathbb{R}^n$ konveksan ako on za bilo koje dvije točke $a, b \in S$ također sadrži i segment određen tim točkama, to jest ako za svaki $a, b \in S$ i za svaki $\lambda \in [0, 1]$ vrijedi*

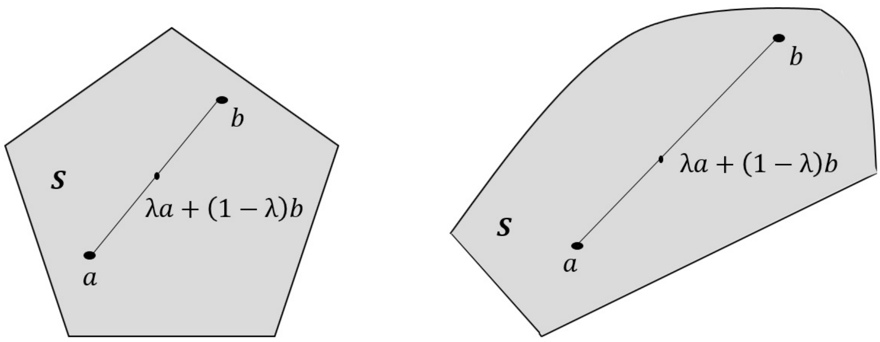
$$\lambda a + (1 - \lambda) b \in S.$$

(ii) Neka je $S \subseteq \mathbb{R}^n$ konveksan skup. Kažemo da je funkcija $f\colon S \to \mathbb{R}$ konveksna ako je graf funkcije ispod svake njegove sekante, to jest ako za svaki $x, y \in S$ i za svaki $\lambda \in [0, 1]$ vrijedi

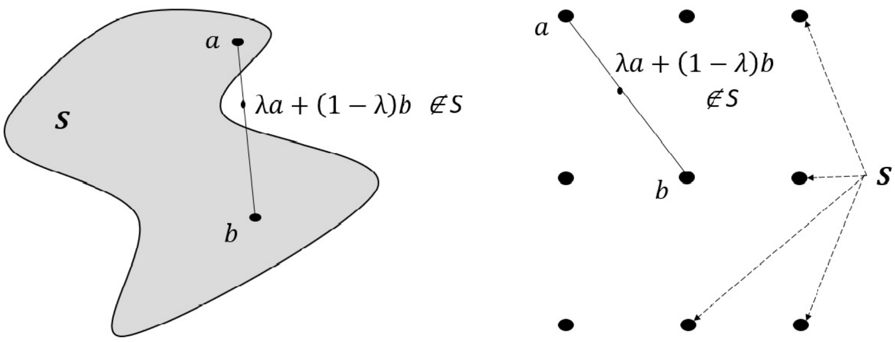
$$f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y).$$

(iii) Ukoliko za sve $x \neq y$ i $\lambda \in \langle 0, 1\rangle$ stoji stroga nejednakost kažemo da je funkcija f strogo konveksna, to jest ako za svaki $x, y \in S$, $x \neq y$ i za svaki $\lambda \in \langle 0, 1\rangle$ vrijedi

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

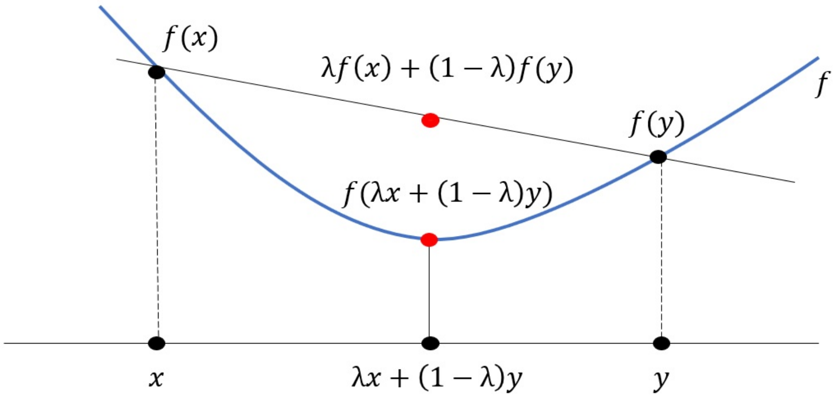


Slika 1.1: Dva konveksna skupa $S \subseteq \mathbb{R}^2$

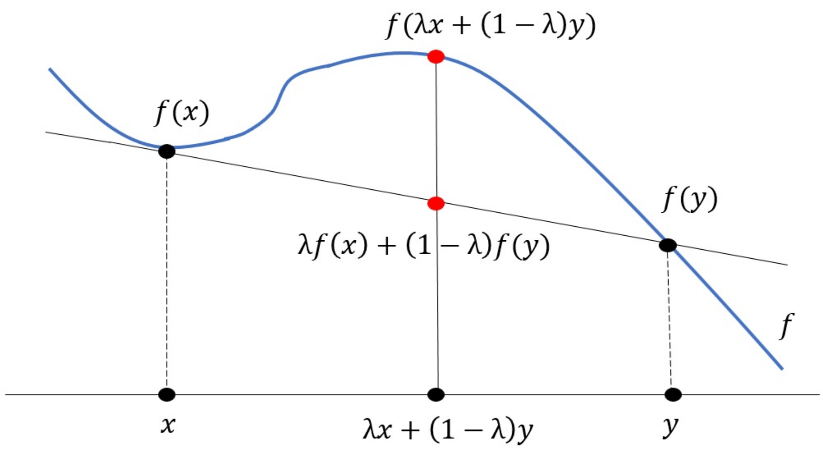


Slika 1.2: Dva nekonveksna skupa $S \subseteq \mathbb{R}^2$

Slike 1.1 i 1.2 prikazuju primjere konveksnih, odnosno nekonveksnih skupova. Uočimo da je u slučaju nekonveksnih skupova dovoljno pronaći barem dvije točke iz skupa za koje njihova spojnica nije sadržana u skupu. Nadalje, Slike 1.3 i 1.4 dodatno objašnjavaju pojam konveksnih i nekonveksnih funkcija.



Slika 1.3: Konveksna funkcija



Slika 1.4: Nekonveksna funkcija

Teorem 1.4 (Fundamentalni teorem globalne optimalnosti)**.** *Neka je $S \subseteq \mathbb{R}^n$ konveksan skup i $f\colon S \to \mathbb{R}$ konveksna funkcija. Tada je svaki lokalni minimum funkcije f na skupu S ujedno i globalni minimum.*

Dokaz. Pretpostavimo suprotno, to jest neka je x^* točka lokalnog minimuma, ali ne i globalnog. Neka je $\bar{x}$ točka globalnog minimuma. Tada iz definicije globalnog minimuma imamo $f(\bar{x}) < f(x^*)$. Uzmimo $\lambda \in \langle 0, 1\rangle$. Zbog konveksnosti skupa S i funkcije f vrijedi

$$\lambda\bar{x} + (1-\lambda)x^* \in S$$

i

$$f(\lambda\bar{x} + (1-\lambda)x^*) \leq \lambda \underbrace{f(\bar{x})}_{< f(x^*)} + (1-\lambda)f(x^*) < f(x^*).$$

Uzimanjem proizvoljno malog $\lambda > 0$ dobivamo kontradikciju s lokalnom optimalnošću točke x^*.

■

U radu će nam biti potreban i sljedeći rezultat o karakterizaciji diferencijabilnih konveksnih funkcija.

Propozicija 1.5. *Neka je f diferencijabilna funkcija na otvorenom, konveksnom skupu S. Tada je f konveksna na S ako i samo ako za svaki $x, y \in S$ vrijedi $f(y) \geq f(x) + \nabla f(x)^T(y - x)$.*

Dokaz. $\boxed{\Rightarrow}$ Neka je f konveksna na S. Uzmimo $x_1, x_2 \in S$ proizvoljne i $\lambda \in \langle 0, 1\rangle$. Tada zbog konveksnosti funkcije f vrijedi

$$\lambda f(x_1) + (1-\lambda)f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2)$$

što je ekvivalentno

$$\lambda f(x_1) - \lambda f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2) - f(x_2).$$

Sređivanjem izraza s desne strane te dijeljenjem nejednadžbe s $\lambda > 0$ dobivamo

$$f(x_1) - f(x_2) \geq \frac{f(x_2 + \lambda(x_1 - x_2)) - f(x_2)}{\lambda}.$$

Uzimanjem limesa kada $\lambda \to 0$ slijedi

$$f(x_1) - f(x_2) \geq \nabla f(x_2)^T(x_1 - x_2),$$

to jest graf konveksne funkcije leži iznad tangencijalne ravnine u svakoj točki $x_2 \in S$. Iz posljednje nejednakosti slijedi tvrdnja.

$\boxed{\Leftarrow}$ Neka za svaki $x, y \in S$ vrijedi $f(y) \geq f(x) + \nabla f(x)^T(y - x)$. Neka su $x_1, x_2 \in S$ proizvoljni i $\lambda \in \langle 0, 1\rangle$. Kako je S konveksan znamo $\lambda x_1 + (1 - \lambda)x_2 \in S$. Uvrstimo sada u prethodnu nejednakost $\lambda x_1 + (1 - \lambda)x_2$ umjesto x te x_1 u prvom, odnosno x_2 u drugom slučaju umjesto y. Sređivanjem se dobije

$$\begin{aligned} f(x_1) &\geq f(\lambda x_1 + (1 - \lambda)x_2) + (1 - \lambda)\nabla f(\lambda x_1 + (1 - \lambda)x_2)^T(x_1 - x_2) \\ f(x_2) &\geq f(\lambda x_1 + (1 - \lambda)x_2) + \lambda \nabla f(\lambda x_1 + (1 - \lambda)x_2)^T(x_2 - x_1). \end{aligned}$$

Množenjem prve nejednažbe s λ i druge s $(1 - \lambda)$ te sumiranjem dobivenih nejednadžbi dobivamo

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2),$$

što po definiciji znači da je f konveksna funkcija.

■

1.2 Koercitivnost i Weierstrassov teorem

Prije nego krenemo razvijati nužne i dovoljne uvjete optimalnosti osvrnimo se na pitanje egzistencije optimalnog rješenja. Osnovni teorem koji govori pod kojim uvjetima minimum postoji jest Weierstrassov teorem.

Teorem 1.6. *Neka je $K \subseteq \mathbb{R}^n$ neprazan, kompaktan[2] skup i $f\colon K \to \mathbb{R}$ neprekidna funkcija. Tada funkcija f postiže globalni minimum i maksimum na K.*

Međutim, najčešće ipak nije zadovoljena pretpostavka na omeđenost skupa i tada se koristi sljedeći oblik Weierstrassovog teorema. Prije samog iskaza, uvodimo novi pojam koercitivne funkcije te karakterizaciju pomoću kompaktnosti.

Definicija 1.7. *Kažemo da je funkcija $f\colon \mathbb{R}^n \to \mathbb{R}$ **koercitivna** ako za svaki niz $(x_n : n \in \mathbb{N}) \subset \mathbb{R}^n$ takav da $\|x_n\| \to +\infty$ vrijedi $f(x_n) \to +\infty$.*

Napomena 1.8. *Koercitivnost funkcije f ekvivalentna je*

$$\forall M > 0, \exists N > 0, \forall x \in S, \quad \|x\| \geq N \Rightarrow f(x) \geq M.$$

Propozicija 1.9. *Neka je $f\colon \mathbb{R}^n \to \mathbb{R}$ neprekidna funkcija. Funkcija f je koercitivna ako i samo ako je nivo skup $\{\, x \mid f(x) \leq \alpha \,\}$ kompaktan za svaki $\alpha \in \mathbb{R}$.*

[2]Skup $K \subset \mathbb{R}^n$ je kompaktan ako je omeđen i zatvoren.

Dokaz. Uočimo da neprekidnost funkcije f implicira zatvorenost skupa $\{ x \mid f(x) \leq \alpha \}$ pa je u dokazu dovoljno analizirati samo omeđenost skupa.
$\boxed{\Rightarrow}$ Neka je f neprekidna, koercitivna funkcija. Pokažimo da je tada skup $\{ x \mid f(x) \leq \alpha \}$ omeđen za svaki $\alpha \in \mathbb{R}$. Pretpostavimo suprotno, to jest neka je $\alpha \in \mathbb{R}$ takav da je skup $S := \{ x \mid f(x) \leq \alpha \}$ neomeđen. Tada postoji niz $(x_n : n \in \mathbb{N}) \subset S$ takav da $\|x_n\| \to +\infty$. Zbog koercitivnosti funkcije f mora vrijediti $f(x_n) \to +\infty$, što je u kontradikciji s činjenicom $f(x_n) \leq \alpha$, za svaki $n \in \mathbb{N}$. Dakle, skup $\{ x \mid f(x) \leq \alpha \}$ je omeđen za svaki $\alpha \in \mathbb{R}$, a time i kompaktan.
$\boxed{\Leftarrow}$ Neka je skup $\{ x \mid f(x) \leq \alpha \}$ omeđen za svaki $\alpha \in \mathbb{R}$. Pokažimo da je tada funkcija f koercitivna. Pretpostavimo suprotno, to jest neka je $(x_n : n \in \mathbb{N}) \subset \mathbb{R}^n$ niz takav da $\|x_n\| \to +\infty$, ali $f(x_n) \not\to +\infty$. Tada postoji podniz $(x_{p_n} : n \in \mathbb{N})$ takav da je $(f(x_{p_n}) : n \in \mathbb{N})$ omeđen odozgo. Dakle, postoji $\alpha \in \mathbb{R}$ takav da je $(x_{p_n} : n \in \mathbb{N}) \in \{ x \mid f(x) \leq \alpha \}$. Budući da je skup $\{ x \mid f(x) \leq \alpha \}$ omeđen za svaki $\alpha \in \mathbb{R}$, dobivamo kontradikciju s pretpostavkom da je svaki podniz niza $(x_n : n \in \mathbb{N})$ neomeđen.

∎

Teorem 1.10 (Weierstrass)**.** *Neka je $S \subseteq \mathbb{R}^n$ neprazan, zatvoren skup i $f \colon \mathbb{R}^n \to \mathbb{R}$ neprekidna, koercitivna funkcija. Tada funkcija f postiže globalni minimum na S.*

Dokaz. Neka je $\alpha \in \mathbb{R}$ takav da je skup $A := \{ x \in S \mid f(x) \leq \alpha \}$ neprazan. Zbog koercitivnosti funkcije f slijedi da je taj skup kompaktan pa po standardnom Weierstrassovom teoremu slijedi da neprekidna funkcija na kompaktu postiže globalni minimum. Očito je da funkcija f na skupu A poprima isti minimum kao i na skupu S.

∎

U slučaju strogo konveksnih funkcija vrijedi i jači rezultat. Stroga konveksnost funkcije cilja osigurava jedinstvenost točke globalnog minimuma i ta tvrdnja je dokazana u sljedećoj propoziciji.

Propozicija 1.11. *Neka je S konveksan skup i $f \colon S \to \mathbb{R}$ strogo konveksna funkcija. Tada za funkciju f postoji najviše jedna točka globalnog minimuma.*

Dokaz. Pretpostavimo da su $x_1, x_2 \in S$, $x_1 \neq x_2$ točke u kojima f poprima globalni minimum na S. Tada iz definicije stroge konveksnosti funkcije f za svaki $\lambda \in \langle 0, 1 \rangle$ vrijedi

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2) = f(x_1),$$

gdje posljednja jednakost proizlazi iz pretpostavke da su x_1 i x_2 točke globalnog minimuma. Zbog konveknosti skupa S vrijedi $\lambda x_1 + (1-\lambda)x_2 \in S$, što znači da smo pronašli cijeli interval točaka za koje je funkcijska vrijednost manja od funkcijske vrijednosti od x_1, a to je u kontradikciji s pretpostavkom globalne optimalnosti točke x_1.

∎

Poglavlje 2

Uvjeti optimalnosti

Na početku poglavlja ćemo se baviti uvjetima optimalnosti za zadaće bezuvjetne optimizacije i za zadaće uvjetne optimizacije u kojima je dopustivi skup zatvoren i konveksan. Zatim prelazimo na općenitiji slučaj i izvodimo geometrijske uvjete optimalnosti kako bismo bolje razumjeli vezu između gradijenta funkcije cilja i dopustivog skupa koja mora biti zadovoljena u svakoj točki lokalnog minimuma. Uz uvođenje dodatnih uvjeta tipa nejednakosti na skup dopustivih točaka, geometrijski će uvjeti implicirati vezu između gradijenta funkcije cilja i aktivnh funkcija ograničenja. Tako dobivamo Fritz Johnove uvjete optimalnosti, koji su ipak slabi i mogu biti zadovoljeni u mnogim točkama koje nemaju veze s optimalnim rješenjima. Međutim, uz dodatnu pretpostavku regularnosti na dopustivi skup dobivamo jače uvjete koji su poznati kao Karush-Kuhn-Tuckerovi. Spomenute uvjete regularnosti ćemo označavati s CQ (engl. constraint qualifications) i navest ćemo neke od poznatijih te dati vezu između njih. Na samom kraju ćemo pokazati da su u zadaćama konveksnog programiranja KKT uvjeti dovoljni za lokalnu, odnosno globalnu optimalnost.

2.1 Bezuvjetna optimizacija

Razvoj uvjeta optimalnosti započinjemo slučajem u kojem je dopustivi skup $S \subseteq \mathbb{R}^n$ otvoren, a funkcija cilja diferencijabilna. Dakle, promatramo sljedeću zadaću optimizacije:

$$\begin{aligned} \min\ & f(x) \\ & x \in S, \end{aligned}$$

pri čemu je $S \subseteq \mathbb{R}^n$ neprazan i otvoren skup te $f\colon S \to \mathbb{R}$ diferencijabilna funkcija.

Teorem 2.1 (Nužni uvjeti prvog reda)**.** *Neka je $f\colon S \to \mathbb{R}$ diferencijabilna u točki $x^* \in S$. Ako je x^* točka lokalnog minimuma funkcije f onda $\nabla f(x^*) = 0$.*

Napomena 2.2. *Gradijent funkcije f u točki x koji označavamo s $\nabla f(x)$ definiran je kao*

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)^T.$$

Uočimo da se tada gornji uvjet svodi na

$$\frac{\partial f(x^*)}{\partial x_j} = 0, \quad j = 1, \ldots, n.$$

Točku x^* koja zadovoljava gornji uvjet zovemo stacionarna točka.
[Za $n = 1$ se tvrdnja teorema svodi na: $x^* \in \mathbb{R}$ točka lokalnog minimuma $\implies f'(x^*) = 0$.]

Dokaz. Pretpostavimo suprotno, to jest neka je x^* točka lokalnog minimuma takva da $\nabla f(x^*) \neq 0$. Neka je $p := -\nabla f(x^*)$. Prema definiciji derivacije funkcije f u smjeru p u točki x^*, imamo

$$f(x^* + \alpha p) = f(x^*) + \alpha \nabla f(x^*)^T p + o(\alpha),$$

gdje je $o\colon \mathbb{R} \to \mathbb{R}$ realna funkcija takva da $o(s)/s \to 0$ kada $s \to 0$. Raspisivanjem dobijemo

$$\begin{aligned} f(x^* + \alpha p) &= f(x^*) + \alpha \nabla f(x^*)^T(-\nabla f(x^*)) + o(\alpha) \\ &= f(x^*) - \alpha \|\nabla f(x^*)\|^2 + o(\alpha) \\ &= f(x^*) - \alpha \left[\|\nabla f(x^*)\|^2 + \frac{o(\alpha)}{\alpha} \right]. \end{aligned}$$

Budući da je po pretpostavci $\|\nabla f(x^*)\| \neq 0$, za dovoljno mali $\alpha > 0$ vrijedi

$$f(x^* + \alpha p) < f(x^*),$$

što je u kontradikciji s činjenicom da je x^* točka lokalnog minimuma.

∎

Obratna implikacija ne mora vrijediti. Kao kontraprimjer uzmimo funkciju $f\colon \mathbb{R} \to \mathbb{R}$ definiranu s $f(x) = x^3$. Tada je $\bar{x} = 0$ stacionarna točka, ali nije točka lokalnog ekstrema. Dokaz prethodnog teorema je važan jer je u njemu konstruiran vektor smjera p takav da mali pomak duž njega iz neke točke x vodi k manjoj funkcijskoj vrijednosti i to nas motivira za uvođenje novog pojma.

Definicija 2.3. *Neka je $f\colon \mathbb{R}^n \to \mathbb{R}$ realna funkcija i $x, p \in \mathbb{R}^n$ vektori. Kažemo da je vektor $p \in \mathbb{R}^n$* ***smjer silaska*** *funkcije f u točki x ako postoji $\delta > 0$ takav da $f(x + \alpha p) < f(x)$, za svaki $\alpha \in \langle 0, \delta\,]$.*

Propozicija 2.4 (Dovoljni uvjeti za smjer silaska)**.** *Neka je $f\colon \mathbb{R}^n \to \mathbb{R}$ diferencijabilna u točki $x \in \mathbb{R}^n$. Ako za vektor $p \in \mathbb{R}^n$ vrijedi $\nabla f(x)^T p < 0$ onda je p smjer silaska funkcije f u točki $x \in \mathbb{R}^n$.*

Dokaz. Iz definicije derivacije funkcije f u smjeru p u točki x, imamo

$$f(x + \alpha p) = f(x) + \alpha \nabla f(x)^T p + o(\alpha).$$

Kako je po pretpostavci $\nabla f(x)^T p < 0$, slijedi da je $f(x + \alpha p) < f(x)$ za dovoljno male $\alpha > 0$.

■

Ukoliko su zadovoljene jače pretpostavke na glatkoću funkcije f, nužne uvjete lokalnog minimuma možemo još preciznije opisati pomoću Hesseove matrice $\nabla^2 f(x)$. Prije nego iskažemo teorem koji govori o tome, definirajmo pojam pozitivno semidefinitnih i definitnih matrica.

Definicija 2.5. *Za kvadratnu matricu $A \in M_n(\mathbb{R})$ kažemo da je:*

(i) ***pozitivno semidefinitna*** *ako je $x^T A x \geq 0$, za svaki $x \in \mathbb{R}^n$,*

(ii) ***pozitivno definitna*** *ako je $x^T A x > 0$, za svaki $x \in \mathbb{R}^n \setminus \{0\}$.*

Teorem 2.6 (Nužni uvjeti drugog reda)**.** *Neka je $f\colon S \to \mathbb{R}$ dva puta diferencijabilna u točki $x^* \in S$. Ako je x^* točka lokalnog minimuma funkcije f onda vrijedi $\nabla f(x^*) = 0$ i $\nabla^2 f(x^*)$ pozitivno semidefinitna matrica.*

[Za $n = 1$ se tvrdnja teorema svodi na: $x^* \in \mathbb{R}$ točka lokalnog minimuma $\Longrightarrow$ $f'(x^*) = 0$ i $f''(x^*) \geq 0$.]

Dokaz. Neka je $p \in \mathbb{R}^n$ proizvoljan vektor. Zbog pretpostavke da je funkcija f dva puta diferencijabilna u točki $x^* \in S$ vrijedi

$$f(x^* + \alpha p) = f(x^*) + \alpha \nabla f(x^*)^T p + \frac{\alpha^2}{2} p^T \nabla^2 f(x^*) p + o(\alpha^2).$$

Budući da je x^* točka lokalnog minimuma, po Teoremu 2.1 vrijedi $\nabla f(x^*) = 0$. Preostaje pokazati $\nabla^2 f(x^*)$ pozitivno semidefinitna matrica. Sređivanjem gornje jednadžbe te dijeljenjem s $\alpha^2 > 0$ dobijemo

$$\frac{f(x^* + \alpha p) - f(x^*)}{\alpha^2} = \frac{1}{2} p^T \nabla^2 f(x^*) p + \frac{o(\alpha^2)}{\alpha^2}.$$

Zbog lokalne optimalnosti točke x^* vrijedi $f(x^* + \alpha p) \geq f(x^*)$ za dovoljno mali $\alpha > 0$ pa je lijeva strana gornje jednadžbe nenegativna. Zbog toga vrijedi

$$\frac{1}{2} p^T \nabla^2 f(x^*) p + \frac{o(\alpha^2)}{\alpha^2} \geq 0.$$

Uzimanjem limesa kada $\alpha \to 0$ slijedi

$$p^T \nabla^2 f(x^*) p \geq 0.$$

Kako je $p \in \mathbb{R}^n$ bio proizvoljan slijedi da je $\nabla^2 f(x^*)$ pozitivno semidefinitna matrica.

∎

Obrat gornjeg teorema ne vrijedi, kontraprimjer je isti kao i za Teorem 2.1. Za funkciju $f\colon \mathbb{R} \to \mathbb{R}$ definiranu s $f(x) = x^3$ je $\bar{x} = 0$ stacionarna točka i vrijedi $f''(x) = 6x$ pa je zadovoljeno $f''(0) \geq 0$, ali ne radi se o točki lokalnog ekstrema.

Sljedeći rezultat govori da uz pretpostavku pozitivne definitnosti Hesseove matrice možemo zaključiti da su stacionarne točke ujedno točke strogog lokalnog minimuma.

Teorem 2.7 (Dovoljni uvjeti drugog reda)**.** *Neka je $f\colon S \to \mathbb{R}$ dva puta diferenijabilna u točki $x^* \in S$. Ako vrijedi $\nabla f(x^*) = 0$ i $\nabla^2 f(x^*)$ pozitivno definitna matrica onda je x^* točka strogog lokalnog minimuma funkcije f.*

[Za $n = 1$ se tvrdnja teorema svodi na: $f'(x^*) = 0$ i $f''(x^*) > 0 \implies x^* \in \mathbb{R}$ točka strogog lokalnog minimuma.]

Dokaz. Neka je $\nabla f(x^*) = 0$ i $\nabla^2 f(x^*)$ pozitivno definitna matrica. Uzmimo proizvoljan vektor $p \in \mathbb{R}^n$, $p \neq 0$. Tada vrijedi

$$\begin{aligned} f(x^* + \alpha p) &= f(x^*) + \alpha \underbrace{\nabla f(x^*)^T p}_{=0} + \frac{\alpha^2}{2} \underbrace{p^T \nabla^2 f(x^*) p}_{>0} + o(\alpha^2) \\ &> f(x^*), \end{aligned}$$

za dovoljno mali $\alpha > 0$. Kako je p bio proizvoljan, gornja nejednakost povlači da je x^* točka strogog lokalnog minimuma funkcije f.

∎

Teoremom 2.1 smo pokazali da je nužan uvjet lokalnog minimuma stacionarnost točke, a u slučaju konveksnih funkcija taj uvjet je i dovoljan za globalnu, odnosno lokalnu optimalnost.

Teorem 2.8 (Nužni i dovoljni uvjeti globalne optimalnosti)**.** *Neka je $f\colon S \to \mathbb{R}$ diferencijabilna konveksna funkcija. Tada je x^* točka globalnog minimuma funkcije f ako i samo ako vrijedi $\nabla f(x^*) = 0$.*

Dokaz. $\boxed{\Rightarrow}$ Neka je x^* točka globalnog minimuma funkcije f. Budući da je točka globalnog minimuma ujedno i točka lokalnog minimuma direktno iz Teorema 2.1 slijedi $\nabla f(x^*) = 0$.
$\boxed{\Leftarrow}$ Neka je $\nabla f(x^*) = 0$. Zbog konveksnosti funkcije f po Propoziciji 1.5 za svaki $y \in \mathbb{R}^n$ vrijedi

$$\begin{aligned} f(y) &\geq f(x^*) + \nabla f(x^*)^T (y - x^*) \\ &= f(x^*), \end{aligned}$$

gdje jednakost u drugom redu vrijedi zbog pretpostavke $\nabla f(x^*) = 0$. Dakle, vrijedi $f(y) \geq f(x^*)$ za svaki $y \in \mathbb{R}^n$, to jest x^* je točka globalnog minimuma funkcije f.

■

2.2 Optimalnost u slučaju konveksnih skupova

U daljnjoj analizi promatramo sljedeću zadaću:

$$\begin{aligned} \min\ & f(x) \\ & x \in S, \end{aligned}$$

gdje je $S \subseteq \mathbb{R}^n$ neprazan, zatvoren i konveksan skup te $f\colon S \to \mathbb{R}$ diferencijabilna funkcija.

Bitna razlika u odnosu na probleme bezuvjetne optimizacije u kojima je $S \subseteq \mathbb{R}^n$ otvoren koje smo promatrali u prethodnom odjeljku jest ta što se u zadaćama u kojima je $S \subset \mathbb{R}^n$ zatvoren može dogoditi da se pomicanjem iz točke x duž nekog vektora p maknemo iz dopustivog skupa S. Dakle, ne vrijedi da je svaki vektor p dopustivi vektor. Možemo li koristiti vektor p kao smjer pomicanja iz točke x će ovisiti o uvjetima koji definiraju dopustivi skup, to jest jesu li oni aktivni u točki x ili ne. Stoga u ovom odjeljku dajemo formalne definicije ovih pojmova kako bismo mogli razviti nužne i dovoljne uvjete optimalnosti bazirane na njima. Započnimo definicijom vektora dopustivog smjera u x, koja kaže da se malim pomicanjem u tom smjeru moramo zadržati unutar dopustivog skupa S.

Definicija 2.9. *Neka je $S \subseteq \mathbb{R}^n$, $x \in S$ i $p \in \mathbb{R}^n$. Kažemo da je p* ***vektor dopustivog smjera*** *u točki x ako postoji $\delta > 0$ takav da $x + \alpha p \in S$, za svaki $\alpha \in [0, \delta]$.*

Prisjetimo se, u prethodnom odjeljku smo pokazali da ukoliko vrijedi $\nabla f(x)^T p < 0$ tada vektor p definira smjer silaska funkcije f u točki x. Sada možemo formulirati nužne uvjete optimalnosti za zadaću uvjetne optimizacije u slučaju konveksnih skupova.

Propozicija 2.10 (Nužni uvjeti prvog reda)**.** *Neka je $S \subseteq \mathbb{R}^n$ i $f\colon S \to \mathbb{R}$ diferencijabilna funkcija.*

(a) Ako je $x^* \in S$ točka lokalnog minimuma funkcije f na skupu S onda za svaki vektor dopustivog smjera p u točki x^* vrijedi $\nabla f(x^*)^T p \geq 0$.

(b) Neka je S konveksan skup. Ako je $x^* \in S$ točka lokalnog minimuma funkcije f na skupu S onda za svaki $x \in S$ vrijedi

$$\nabla f(x^*)^T (x - x^*) \geq 0. \tag{2}$$

Dokaz. (a) Tvrdnju dokazujemo obratom po kontrapoziciji. Neka je p vektor dopustivog smjera takav da vrijedi $\nabla f(x^*)^T p < 0$. Tada slijedi da je p vektor silaska, to jest vrijedi $f(x^* + \alpha p) < f(x^*)$ za dovoljno mali $\alpha > 0$. Dakle, x^* nije točka lokalnog minimuma.
(b) Budući da je S po pretpostavci konveksan skup, svaki vektor dopustivog smjera p možemo zapisati kao $\alpha(x - x^*)$, za neki $x \in S$ i $\alpha > 0$. Tada tvrdnja slijedi iz (a) dijela.

■

Ranije smo pokazali (Teorem 2.1) da je za zadaće bezuvjetne optimizacije u kojima je funkcija cilja f diferencijabilna nužan uvjet optimalnosti $\nabla f(x^*) = 0$. Uočimo da taj uvjet proizlazi iz nejednakosti (2) za slučaj otvorenog skupa $S \subseteq \mathbb{R}^n$, jer je jedini način da nejednakost bude zadovoljena za svaki $x \in S$ ako vrijedi $\nabla f(x^*) = 0$. Kao i u slučaju bezuvjetne optimizacije, vektor $x^* \in S$ koji zadovoljava nejednakost (2) ćemo zvati stacionarna točka.

U nastavku donosimo rezultat za zadaću konveksnog programiranja, to jest minimizacije konveksne funkcije f na konveksnom skupu S. Primijetimo da se radi o poopćenju Teorema 2.8.

Teorem 2.11 (Nužni i dovoljni uvjeti globalne optimalnosti)**.** *Neka je $S \subseteq \mathbb{R}^n$ neprazan i konveksan skup te $f: S \to \mathbb{R}$ diferencijabilna konveksna funkcija. Tada je x^* točka globalnog minimuma funkcije f na skupu S ako i samo ako vrijedi $\nabla f(x^*)^T (x - x^*) \geq 0$, za svaki $x \in S$.*

Dokaz. $\boxed{\Rightarrow}$ Neka je x^* točka globalnog minimuma funkcije f na S. Budući da je točka globalnog minimuma ujedno i točka lokalnog minimuma tvrdnja direktno slijedi iz (b) dijela Propozcije 2.10.
$\boxed{\Leftarrow}$ Neka je $\nabla f(x^*)^T (x - x^*) \geq 0$, za svaki $x \in S$. Kako je f konveksna, po Propoziciji 1.5 za svaki $y \in S$ vrijedi

$$\begin{aligned} f(y) &\geq f(x^*) + \nabla f(x^*)^T (y - x^*) \\ &\geq f(x^*), \end{aligned}$$

gdje nejednakost u drugom redu vrijedi zbog pretpostavke $\nabla f(x^*)^T (x - x^*) \geq 0$. Dakle, vrijedi $f(x^*) \leq f(y)$, za svaki $y \in S$, to jest x^* je točka globalnog minimuma funkcije f na skupu S.

■

Na kraju ovog odjeljka uvodimo pojam normalnog konusa čija će relevantnost biti istaknuta kasnije u radu. Ujedno, u terminima istog, iskazat ćemo ekvivalentnu definiciju stacionarne točke.

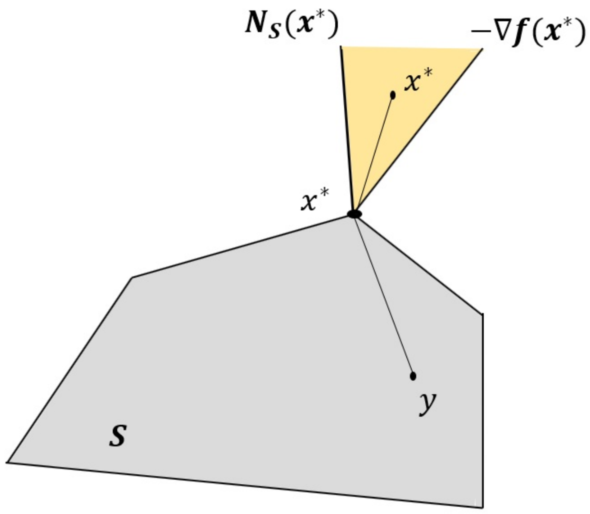
Definicija 2.12. *Neka je $S \subseteq \mathbb{R}^n$ zatvoren i konveksan skup.* ***Normalni konus*** *na S u točki $x \in S$ definira se kao:*

$$N_S(x) := \{\, p \in \mathbb{R}^n \mid p^T(y - x) \leq 0, y \in S \,\}.$$

Sada kažemo da je x^* stacionarna točka ako vrijedi

$$-\nabla f(x^*) \in N_S(x^*). \tag{3}$$

Očito je da je ovaj uvjet ekvivalentan uvjetu (2) i po Teoremu 2.11 oba uvjeta su nužna i dovoljna za globalnu optimalnost točke x^* za konveksne funkcije f. Geometrijski gledano, uvjet (3) kaže da je kut kojeg zatvaraju negativni gradijent funkcije cilja i bilo koji dopustivi smjer veći ili jednak 90° (vidi Slika 2.1). To je ekvivalentno činjenici da u točki x^* ne postoji dopustivi smjer silaska.



Slika 2.1: Karakterizacija stacionarne točke u terminima normalnog konusa

2.3 Geometrijski uvjeti optimalnosti

Od sada pa nadalje analiziramo sljedeću zadaću nelinearnog programiranja:

$$\begin{aligned} &\min f(x) \\ &x \in S, \end{aligned} \tag{4}$$

gdje je $S \subseteq \mathbb{R}^n$ neprazan zatvoren skup i $f\colon S \to \mathbb{R}$ diferencijabilna funkcija. Primijetimo da smo u prethodnom odjeljku već analizirali jednostavniji slučaj ove zadaće u kojem je S konveksan i f diferencijabilna na S i pokazali smo da vrijedi:

Ako je x^* točka lokalnog minimuma onda je x^* stacionarna točka.

Budući da nemamo nikakvih ograničenja tipa jednakosti i nejednakosti na dopustivi skup S uvjete optimalnosti ćemo temeljiti na geometriji. Intuitivno, ako je točka $x^* \in S$ lokalni minimum funkcije f na skupu S, onda se vrijednost funkcije f niti u jednom smjeru koji je dopustiv ne smije smanjivati. Kako bismo matematički opisali značenje prethodne rečenice uvodimo prirodnu definiciju skupa zraka koje proizvoljno blizu točke x^* pripadaju skupu S.

Definicija 2.13. *Neka je $S \subseteq \mathbb{R}^n$ neprazan zatvoren skup.* ***Konus dopustivih smjerova*** *za S u $x \in \mathbb{R}^n$ (engl.radial cone) definira se kao:*

$$R_S(x) := \{\, p \in R^n \mid \exists \delta > 0 \textit{ takav da } x + \alpha p \in S, \alpha \in [\,0, \delta]\,\}.$$

Dakle, $R_S(x)$ je konus koji sadrži sve vektore dopustivih smjerova u smislu Definicije 2.9. Vektor p je element konusa $R_S(x)$ ako dopustivi skup S sadrži netrivijalni dio zrake $x+\alpha p, \alpha \geq 0$. Nažalost, ovaj konus je premali da bi opisao uvjete optimalnosti za probleme nelinearnog programiranja. Primjerice, za $S = \{\, x \in \mathbb{R}^2 \mid \|x\| = 1\,\}$ vrijedi $R_S(x) = \emptyset$ za svaki $x \in S$. Također, za $S = \{\, x \in \mathbb{R}^2 \mid x_2 = x_1^2\,\}$ vrijedi $R_S(x) = \emptyset$ za svaki $x \in S$. Stoga definiramo kompliciraniji, ali veći skup s lijepim svojstvima koji će nam biti od pomoći.

Definicija 2.14. *Neka je $S \subseteq \mathbb{R}^n$ neprazan zatvoren skup.* ***Tangencijalni konus*** *za S u $x \in \mathbb{R}^n$ definira se kao:*

$$\begin{aligned} T_S(x) := \{\, p \in R^n \mid \exists\, (x_k : k \in \mathbb{N}) \subset S, (\lambda_k : k \in \mathbb{N}) \subset (\,0, \infty) \textit{ takvi da} \\ \lim_{k\to\infty} x_k = x,\ \lim_{k\to\infty} \lambda_k(\,x_k - x) = p\,\}. \end{aligned}$$

Gornja definicija kaže da pri provjeri pripada li vektor p skupu $T_S(x)$, moramo provjeriti postoji li niz dopustivih točaka $(x_k : k \in \mathbb{N})$ u S koji konvergiraju danom $x \in R^n$ takav da je p tangencijalni vektor na niz $(x_k : k \in \mathbb{N})$ u točki x. Dakle, $T_S(x)$ sadrži sve smjerove u kojima se možemo asimptotski približiti točki x unutar skupa S.

Promatrajući isti skup kao i ranije, $S = \{\, x \in \mathbb{R}^2 \mid x_2 = x_1^2\,\}$ vidimo da je $T_S(0) = \{\, p \in \mathbb{R}^2 \mid p_2 = 0\,\}$.

Propozicija 2.15. *Tangencijalni konus je zatvoren konus. Također vrijedi inkluzija*

$$cl\, R_S(x) \subseteq T_S(x),\ x \in R^n.$$

Dokaz. Trivijalno je da je tangencijalni konus zaista konus. Pokažimo da je zatvoren. Neka je $(p_k : k \in \mathbb{N})$ niz u $T_S(x)$ i pretpostavimo da $p_k \to p$. Za svaki $p_k \in T_S(x)$ postoje $x_k \in S$ i $\lambda_k > 0$ takvi da je $\|x_k - x\| < k^{-1}$ i $\|\lambda_k(x_k - x) - p_k\| < k^{-1}$. Očito je da $x_k \to x$ i zbog nejednakosti trokuta vrijedi $\|\lambda_k(x_k - x) - p_k\| \leq \|\lambda_k(x_k - x) - p_k\| + \|p_k - p\|$. Kako desna strana teži k nuli za velike $k \in N$ slijedi da je $p \in T_S(x)$, to jest $T_S(x)$ je zatvoren skup.
Kako smo dokazali da je $T_S(x)$ zatvoren, za dokaz inkluzije je dovoljno pokazati da vrijedi $R_S(x) \subseteq T_S(x)$. Neka je $p \in R_S(x)$ proizvoljan. Tada za dovoljno veliki $k \in \mathbb{N}$ vrijedi $x + k^{-1}p \in S$. Uzmimo $x_k = x + k^{-1}p$ i $\lambda_k = k$. Tada iz definicije tangencijalnog konusa slijedi tvrdnja.

∎

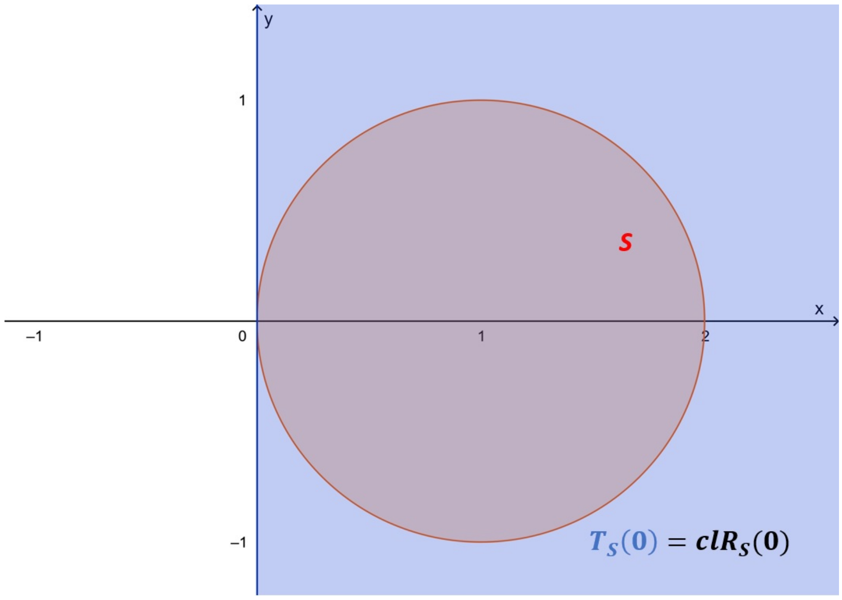
Sada ćemo na nekoliko primjera ilustrirati odnos ova dva konusa. Dopustivi skupovi koje definiramo u primjerima će biti baza za daljnju analizu u radu jer ćemo na njima objašnjavati novouvedene pojmove.

Primjer 2.16. *Neka je* $S := \{ x \in \mathbb{R}^2 \mid -x_1 \leq 0;\ (x_1 - 1)^2 + x_2^2 \leq 1 \}$*. Tada je* $R_S(0) = \{ p \in \mathbb{R}^2 \mid p_1 > 0 \}$ *i* $T_S(0) = \{ p \in \mathbb{R}^2 \mid p_1 \geq 0 \}$*, to jest u ovom primjeru vrijedi* $T_S(0) = cl\, R_S(0)$ *(vidi Slika 2.2).*

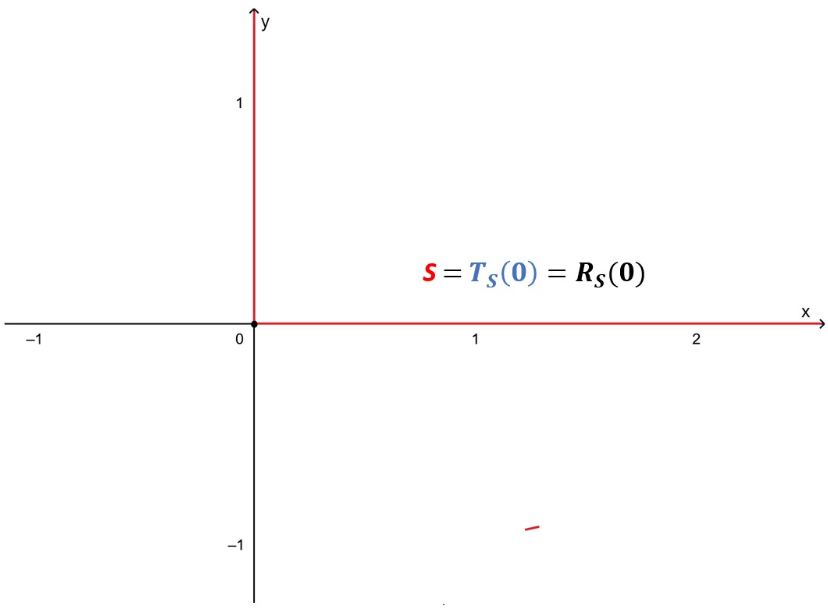
Primjer 2.17. *Neka je* $S := \{ x \in \mathbb{R}^2 \mid -x_1 \leq 0;\ -x_2 \leq 0; x_1x_2 \leq 0 \}$*. U ovom slučaju je* S *nekonveksan konus i vrijedi* $R_S(0) = T_S(0) = S$ *(vidi Slika 2.3).*

Primjer 2.18. *Neka je* $S := \{ x \in \mathbb{R}^2 \mid -x_1^3 + x_2 \leq 0;\ x_1^5 - x_2 \leq 0;\ -x_2 \leq 0 \}$*. Tada je* $R_S(0) = \emptyset$ *i* $T_S(0) = \{ p \in \mathbb{R}^2 \mid p_1 \geq 0;\ p_2 = 0 \}$ *(vidi Slika 2.4).*

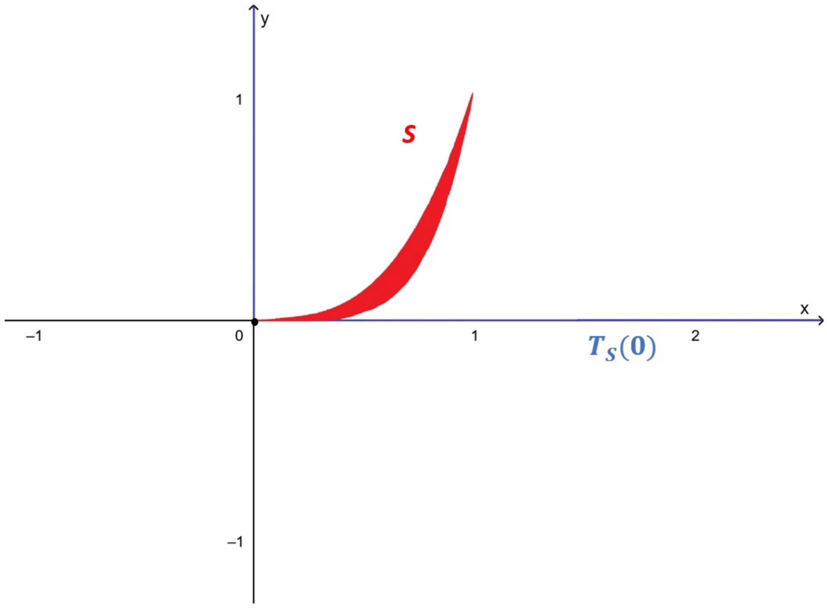
Primjer 2.19. *Neka je* $S := \{ x \in \mathbb{R}^2 \mid -x_2 \leq 0;\ (x_1 - 1)^2 + x_2^2 = 1 \}$*. Tada je* $R_S(0) = \emptyset$ *i* $T_S(0) = \{ p \in \mathbb{R}^2 \mid p_1 = 0;\ p_2 \geq 0 \}$ *(vidi Slika 2.5).*



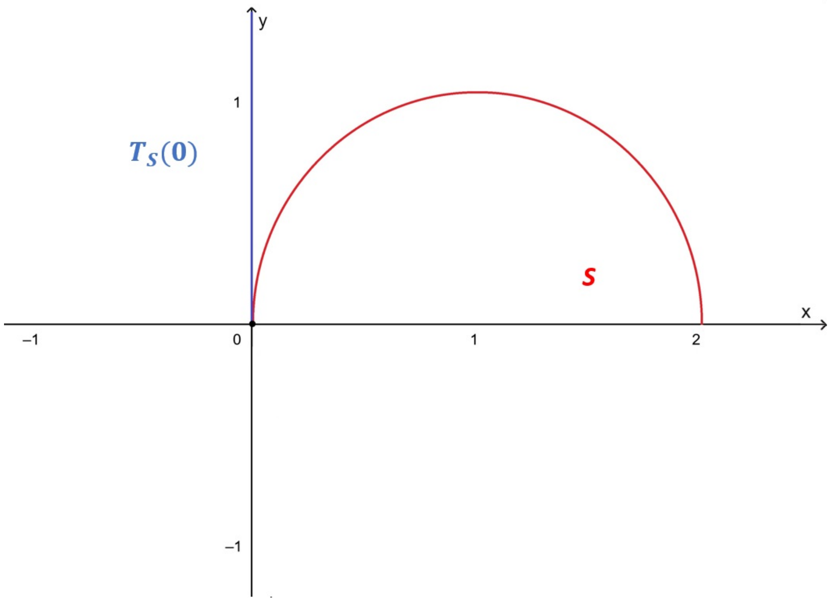
Slika 2.2: Dopustivi skup S i tangencijalni konus $T_S(0)$



Slika 2.3: Dopustivi skup S, tangencijalni konus $T_S(0)$ i konus dopustivih smjerova $R_S(0)$



Slika 2.4: Dopustivi skup S i tangencijalni konus $T_S(0)$, $R_S(0) = \emptyset$



Slika 2.5: Dopustivi skup S i tangencijalni konus $T_S(0)$, $R_S(0) = \emptyset$

Znamo da se vrijednost funkcije f smanjuje duž bilo kojeg smjera silaska (Definicija 2.3) i pokazali smo da je dovoljno da vektor $p \in \mathbb{R}^n$ zadovoljava uvjet $\nabla f(x^*)^T p < 0$ da bi bio vektor smjera silaska funkcije f u točki $x^* \in \mathbb{R}^n$. Budući da je taj uvjet vrlo lako provjeriti koristit ćemo ga za daljnji razvoj uvjeta optimalnosti. Zato definiramo konus takvih smjerova:

$$\mathring{F}(x^*) := \{\, p \in \mathbb{R}^n \mid \nabla f(x^*)^T p < 0 \,\}.$$

Primjetimo da je ovako definirani konus prazan ako je $\nabla f(x^*) = 0$. Sada imamo svu potrebnu notaciju kako bismo iskazali i dokazali glavni teorem ovog odjeljka.

Teorem 2.20 (Nužni geometrijski uvjeti optimalnosti)**.** *Ako je $x^* \in S$ točka lokalnog minimuma onda vrijedi $\mathring{F}(x^*) \cap T_S(x^*) = \emptyset$.*

Dokaz. Neka je $p \in T_S(x^*)$ proizvoljan. Tada postoje nizovi $(x_k : k \in \mathbb{N})$ u S i $(\lambda_k : k \in \mathbb{N})$ u $\langle 0, +\infty\rangle$ takvi da je $\lim_{k\to\infty} x_k = x^*$ i $lim_{k\to\infty}\lambda_k(x_k - x^*) = p$. Prema definiciji gradijenta funkcije f imamo

$$f(x_k) - f(x^*) = \nabla f(x^*)^T(x_k - x^*) + o(\|x_k - x^*\|) \geq 0,$$

gdje posljednja nejednakost vrijedi za dovoljno veliki k zbog lokalne optimalnosti točke x^*. Množenjem jednadžbe s $\lambda_k > 0$ i puštanjem limesa dobivamo sljedeće

$$\begin{aligned} 0 &\leq \lim_{k\to\infty}\left[\lambda_k \nabla f(x^*)^T(x_k - x^*) + \|\lambda_k(x_k - x^*)\| \frac{o(\|x_k - x^*\|)}{\|x_k - x^*\|}\right] \\ &= \nabla f(x^*)^T p + \|p\| \cdot 0, \end{aligned}$$

što povlači $p \notin \mathring{F}(x^*)$, to jest vrijedi $\mathring{F}(x^*) \cap T_S(x^*) = \emptyset$.

■

Prethodni teorem daje elegantan kriterij za pronalazak kandidata za točku lokalnog minimuma zadaće (4), ali problem je što je skup $T_S(x)$ u praksi vrlo teško (gotovo nemoguće) izračunati za općenite skupove S. Jedan način na koji možemo izbjeći računanje tangencijalnog konusa jest da ga zamijenimo nekim manjim konusom i to će nas dovesti do Fritz-Johnovih uvjeta. Uvođenjem manjih konusa dobivamo slabije uvjete optimalnosti. Na primjer, tangencijalni konus možemo zamijeniti konusom dopustivih smjerova jer se kao jednostavna posljedica geometrijskih uvjeta optimalnosti i Propozicije 2.15 ističe sljedeći rezultat:

Ako je $x^* \in S$ točka lokalnog minimuma funkcije f onda vrijedi $\mathring{F}(x^*) \cap R_S(x^*) = \emptyset$.

Kako bismo pokazali koliko je ta tvrdnja slaba promotrimo ponovno primjer u kojem je $S = \{\, x \in \mathbb{R}^2 \mid x_2 = x_1^2 \,\}$. Budući da je $R_S(x) = \emptyset$, nužan uvjet optimalnosti $\mathring{F}(x^*) \cap R_S(x^*) = \emptyset$ vrijedi za svaki dopustivi $x \in S$.

Drugi način na koji možemo izbjeći računanje tangencijalnog konusa jest uvođenje dodatnog zahtjeva regularnosti koji vodi do Karush-Kuhn-Tuckerovih uvjeta optimalnosti. Glavni nedostatak ovakvog pristupa jest što su KKT uvjeti, iako jednako snažni kao geometrijski, manje općeniti i nisu primjenjivi na iregularne probleme.

2.4 Fritz Johnovi uvjeti

U ovom odjeljku uvodimo dodatne uvjete tipa nejednakosti na dopustivi skup S koji će nam olakšati računanje novih konusa koji dobro aproksimiraju $T_S(x^*)$.

Dakle, pretpostavljamo da je dopustivi skup S definiran sustavom nejednadžbi na sljedeći način:

$$S := \{ x \in R^n \mid g_i(x) \leq 0, \ i = 1, \ldots, m \},$$

pri čemu su funkcije g_i neprekidno diferencijabilne na $\mathbb{R}^n$, $i = 1, \ldots, m$. Uvijek možemo pretpostaviti ovakvu strukturu zapisa jer bilo koji uvjet tipa jednakosti $h(x) = 0$ ekvivalentno možemo zapisati u obliku nejednakosti kao $h(x) \leq 0, h(x) \geq 0$.

Za početak definirajmo skup

$$I(x) := \{ i \in \{ 1, \ldots, m \} \mid g_i(x) = 0 \}$$

kao skup indeksa aktivnih funkcija ograničenja, to jest onih koje su u točki x zadovoljene u obliku jednakosti. Sa $|I(x)|$ ćemo označavati kardinalitet tog skupa.

Da bismo mogli aproksimirati tangencijalni konus $T_S(x)$ definiramo konuse koji se odnose na aktivne funkcije ograničenja u danoj točki:

$$\mathring{G}(x) := \{ p \in \mathbb{R}^n \mid \nabla g_i(x)^T p < 0, i \in I(x) \},$$

i

$$G(x) := \{ p \in \mathbb{R}^n \mid \nabla g_i(x)^T p \leq 0, i \in I(x) \}.$$

Primjetimo da $\mathring{G}$ sadrži sve smjerove silaska za aktivne funkcije ograničenja, a G sve one koji nisu smjerovi rasta aktivnih funkcija ograničenja. Sljedeća propozicija opisuje odnos skupova $\mathring{G}, G, R_S$ i T_S. Preciznije, govori da su skupovi $\mathring{G}$ i G dobre aproksimacije za T_S.

Propozicija 2.21. *Za svaki $x \in \mathbb{R}^n$ vrijedi $\mathring{G}(x) \subseteq R_S(x) \subseteq T_S(x) \subseteq G(x)$.*

Dokaz. Dokaz provodimo u tri koraka.

(i) Pokažimo najprije inkluziju $\mathring{G}(x) \subseteq R_S(x)$. Neka je $p \in \mathring{G}(x)$ proizvoljan. Za svaki $i \notin I(x)$ funkcija g_i je neprekidna i vrijedi $g_i(x) < 0$. Stoga je $g_i(x + \delta p) < 0$ za dovoljno male $\delta > 0$. Po Propoziji 2.4 slijedi da je p smjer silaska funkcije g_i u točki x, $i \in I(x)$, što znači da je $g_i(x + \delta p) < g_i(x) = 0$ za svaki takav i te dovoljno mali $\delta > 0$. Dakle, $p \in R_S(x)$, to jest vrijedi $\mathring{G}(x) \subseteq R_S(x)$.

(ii) Sada pokažimo da vrijedi $T_S(x) \subseteq G(x)$. Neka je sada $p \in T_S(x)$, to jest pretpostavimo da postoje nizovi $(x_k : k \in \mathbb{N})$ u S i $(\lambda_k : k \in \mathbb{N})$ u $\langle 0, +\infty\rangle$ takvi da je $\lim_{k\to\infty} x_k = x$ i $\lim_{k\to\infty} \lambda_k(x_k - x) = p$. Analogno kao i u dokazu Teorema 2.20, prema definiciji gradijenta funkcije $g_i, i \in I(x)$, imamo

$$0 \geq g_i(x_k) = g_i(x_k) - g_i(x) = \nabla g_i(x)^T(x_k - x) + o(\|x_k - x\|),$$

gdje prva nejednakost vrijedi zbog dopustivosti x_k. Množenjem nejednadžbe s $\lambda_k > 0$ i puštanjem limesa, za svaki $i \in I(x)$ dobivamo

$$\begin{aligned} 0 &\geq \lim_{k\to\infty} \left[\lambda_k \nabla g_i(x)^T(x_k - x) + \|\lambda_k(x_k - x)\| \frac{o(\|x_k - x\|)}{\|x_k - x\|} \right] \\ &= \nabla g_i(x)^T p + \|p\| \cdot 0, \end{aligned}$$

odakle slijedi $p \in G(x)$.

(iii) Pokažimo još $R_S(x) \subseteq T_S(x)$. Po Propoziciji 2.15 znamo da vrijedi $cl\, R_S(x) \subseteq T_S(x)$, za svaki $x \in R^n$, iz čega zbog definicije zatvarača[1] direktno slijedi $R_S(x) \subseteq T_S(x)$.

Konačno, iz (i), (ii) i (iii) slijedi tvrdnja propozicije.

■

Na sljedećim primjerima ćemo ilustrirati prethodnu propoziciju.

Primjer 2.22 (Nastavak primjera 2.16)**.** *Skup S je određen s dva uvjeta tipa nejednakosti: $g_1(x) := -x_1 \leq 0$ i $g_2(x) := (x_1 - 1)^2 + x_2^2 - 1 \leq 0$. Odredimo $\mathring{G}(0)$ i $G(0)$. Obje funkcije ograničenja su aktivne u zadanoj točki, to jest $g_1(0) = 0$ i $g_2(0) = 0$ pa je skup aktivnih indeksa $I(0) = \{1, 2\}$. Lako se provjeri da je $\nabla g_1(0) = (-1, 0)^T$ i $\nabla g_2(0) = (-2, 0)^T$. Slijedi $\mathring{G}(0) = \{x \in \mathbb{R}^2 \mid x_1 > 0\} = R_S(0)$ i $G(0) = \{x \in \mathbb{R}^2 \mid x_1 \geq 0\} = T_S(0)$. Dakle, za skupove vrijedi inkluzija $\mathring{G}(0) = R_S(0) \subseteq T_S(0) = G(0)$.*

Primjer 2.23 (Nastavak primjera 2.17)**.** *Skup S je određen s tri uvjeta tipa nejednakosti: $g_1(x) := -x_1 \leq 0$, $g_2(x) := -x_2 \leq 0$, $g_3(x) := x_1 x_2 \leq 0$. Sve tri funkcije ograničenja su aktivne u točki $x = 0$ pa je $I(0) = \{1, 2, 3\}$. Lako se provjeri da je $\nabla g_1(0) = (-1, 0)^T$, $\nabla g_2(0) = (0, -1)^T$ i $\nabla g_3(0) = (0, 0)^T$. Slijedi $\mathring{G}(0) = \emptyset \subsetneq R_S(0)$ i $G(0) = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0\} \supsetneq T_S(0)$. Dakle, vrijedi $\mathring{G}(0) \subsetneq R_S(0) = T_S(0) \subsetneq G(0)$.*

[1]Prisjetimo se, zatvarač skupa $A \subseteq \mathbb{R}^n$ je presjek svih zatvorenih skupova što sadrže A, to jest to je najmanji zatvoreni skup koji ga sadrži.

Primjer 2.24 (Nastavak primjera 2.18)**.** *Skup S je određen s tri uvjeta tipa nejednakosti: $g_1(x) := -x_1^3 + x_2 \leq 0$, $g_2(x) := x_1^5 - x_2 \leq 0$, $g_3(x) := -x_2 \leq 0$ i sve tri funkcije ograničenja su aktivne u $x = 0$ pa je $I(0) = \{1, 2, 3\}$. Lako se provjeri da je $\nabla g_1(0) = (0, 1)^T$, $g_2(0) = (0, -1)^T$ i $\nabla g_3(0) = (0, -1)^T$. Slijedi $\mathring{G}(0) = \emptyset = R_S(0)$ i $G(0) = \{ x \in \mathbb{R}^2 \mid x_2 = 0 \} \supsetneq T_S(0)$. Dakle, vrijedi $\mathring{G}(0) = R_S(0) \subsetneq T_S(0) \subsetneq G(0)$.*

Primjer 2.25 (Nastavak primjera 2.19)**.** *Skup S je određen s jednim uvjetom tipa nejednakosti $g_1(x) := -x_2 \leq 0$ i jednim uvjetom tipa jednakosti $h_1(x) := (x_1 - 1)^2 + x_2^2 - 1 = 0$. Drugi uvjet možemo ekvivalentno zapisati pomoću dva uvjeta tipa nejednakosti: $g_2(x) := h_1(x) \leq 0$ i $g_3(x) := -h_1(x) \leq 0$. Time smo dobili da skup S definiraju tri funkcije ograničenja $g_i, i = 1, 2, 3$ i sve tri su aktivne u $x = 0$ pa je $I(0) = \{1, 2, 3\}$. Lako se provjeri da je $\nabla g_1(0) = (0, -1)^T$, $g_2(0) = (-2, 0)^T$ i $\nabla g_3(0) = (2, 0)^T$. Slijedi $\mathring{G}(0) = \emptyset = R_S(0)$ i $G(0) = \{ x \in \mathbb{R}^2 \mid x_1 = 0, x_2 \geqq 0 \} = T_S(0)$. Konačno, imamo $\mathring{G}(0) = R_S(0) \subsetneq T_S(0) = G(0)$.*

Sljedeći važan nužan uvjet optimalnosti, u literaturi poznat kao Fritz Johnov[2], dobivamo kada u geometrijskim uvjetima tangencijalni konus zamijenimo donjom aproksimacijom $\mathring{G}(x)$. Time dobivamo sljedeću tvrdnju:

Ako je x^* točka lokalnog minimuma onda vrijedi $\mathring{G}(x^*) \cap \mathring{F}(x^*) = \emptyset$.

Ovaj uvjet izgleda prilično apstraktno, ali moguće ga je zapisati u praktičnijoj formi i u tome će nam pomoći Farkaseva lema. Gornja jednadžba kaže da za fiksni x^* linearni sustav nejednadžbi nema rješenja. Srećom, Farkaseva lema nam omogućuje da sustav koji nema rješenje zapišemo kao sustav koji ima rješenje. Zato ćemo prije nego krenemo dalje, iskazati i dokazati taj važan rezultat.

Propozicija 2.26 (Farkas)**.** *Neka je $A \in M_{m,n}(\mathbb{R})$, $m \leq n$ i $b \in \mathbb{R}^m$. Tada od sljedeća dva sustava točno jedan ima rješenje:*

(i) $Ax = b$, $x \geq 0$

(ii) $A^T y \leq 0$, $b^T y > 0$.

Dokaz. Prvo ćemo pokazati da oba sustava ne mogu istovremeno imati rješenje, a zatim ćemo pokazati da drugi sustav ima rješenje ako prvi nema.
Kada bi oba sustava imala rješenje, to jest kada bi postojali $\bar{x} \geq 0$ i $\bar{y} \in \mathbb{R}^m$ takvi da vrijede nejednakosti iz iskaza propozicije, onda bi bilo $\bar{x}^T A^T \bar{y} = b^T \bar{y} > 0$. Kako je $\bar{x} \geq 0$ mora vrijediti $A^T \bar{y} > 0$, što je u kontradikciji sa sustavom (ii).
Pretpostavimo sada da prvi sustav nema rješenje, to jest da $b \notin C$, pri čemu je $C := \{ z \in \mathbb{R}^m \mid Ax = z,\ x \geq 0 \}$ konačnogenerirani konus. Budući da je konačnogenerirani konus

[2]Fritz John (Berlin, 14. lipanj 1910. – New York, 10. veljače 1994.), njemački matematičar

konveksan i zatvoren skup, po teoremu separacije[3] postoje $y \in \mathbb{R}^m, y \neq 0$ i $\alpha \in \mathbb{R}$ takvi da je $b^T y > \alpha$ i $z^T y \leq \alpha$, za svaki $z \in C$, to jest

$$x^T A^T y \leq \alpha, \quad x \geq 0. \tag{5}$$

Specijalno, za $x = 0$ dobivamo $\alpha \geq 0$ i $b^T y > 0$. Preostaje pokazati da je $A^T y \leq 0$, to jest da je $a_i^T y \leq 0$ za sve $i = 1, \ldots, n$, pri čemu su $a_1, \ldots, a_n$ stupci matrice A. Iz nejednakosti (5) za $x = e_i$ direktno zaključujemo $a_i^T y \leq 0$ za sve $i = 1, \ldots, n$.

■

Napomena 2.27. *Sustav (ii) iz prethodne propozije možemo zapisati kao $A^T y \geq 0$, $b^T y < 0$.*

Sada smo spremni iskazati teorem koji opisuje Fritz Johnove uvjete optimalnosti u povoljnijoj formi.

Teorem 2.28 (Fritz Johnovi nužni uvjeti)**.** *Ako je $x^* \in S$ točka lokalnog minimuma onda postoje $\mu_0 \in \mathbb{R}$ i $\mu \in \mathbb{R}^m$ takvi da vrijedi*

$$\mu_0 \nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0, \tag{6a}$$

$$\mu_i g_i(x^*) = 0, \quad i = 1, \ldots, m, \tag{6b}$$

$$\mu_0, \mu_i \geq 0, \quad i = 1, \ldots, m, \tag{6c}$$

$$(\mu_0, \mu^T)^T \neq 0. \tag{6d}$$

Dokaz. Iz Propozicije 2.21 i geometrijskih uvjeta optimalnosti slijedi da ne postoji vektor $p \in \mathbb{R}^n$ takav da $\nabla f(x^*)^T p < 0$ i $\nabla g_i(x^*)^T p < 0, i \in I(x^*)$. Definirajmo matricu A na sljedeći način:

$$A = \begin{pmatrix} \nabla f(x^*) & \nabla g_i(x^*) \end{pmatrix}, \quad i \in I(x^*).$$

Tada sustav $A^T p < 0$ nema rješenja. Po Farkasevoj lemi (Propozicija 2.26) postoji vektor $\lambda \in \mathbb{R}^{1+|I(x^*)|} \neq 0$ takav da $\lambda \geq 0$ i $A\lambda = 0$. Definirajmo sada $(\mu_0, \mu_{I(x^*)}^T)^T := \lambda$ i stavimo $\mu_i = 0$ za $i \notin I(x^*)$. Lako se provjeri da ovako definirani μ_0 i μ zadovoljavaju uvjete (6a) – (6d).

■

Rješenje (μ_0, μ) sustava (6) zovemo Lagrangeov multiplikator. Uvjet (6b) se naziva uvjet komplementarnosti. Sve uvjete zajedno, (6a) – (6d) ćemo kraće zvati Fritz Johnov sustav, a točku koja ih zadovoljava Fritz Johnova točka. Primjetimo da svaki multiplikator (osim μ_0) vežemo s odgovarajućom funkcijom ograničenja g_i, to jest odgovara jednom

[3] Teorem separacije: Neka je $C \subseteq \mathbb{R}^n$ zatvoren i konveksan skup i $y \notin C$. Tada postoje vektor $q \neq 0$ i skalar $\alpha \in \mathbb{R}$ takvi da $q^T y > \alpha$ i $q^T x \leq \alpha$, za svaki $x \in C$.

uvjetu u definiciji skupa S. Iz uvjeta komplementarnosti slijedi da multiplikatori μ_i koji odgovaraju neaktivnim funkcijama ograničenja, to jest za $i \notin I(x^*)$, moraju biti jednaki nuli. Općenito, Lagrangeov multiplikator μ_i govori koliko je optimalno rješenje osjetljivo na male promjene u uvjetima g_i.

Primjer 2.29. *Promotimo diferencijabilnu funkciju $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiranu s $f(x) = x_1$. Vrijedi $\nabla f = (1,0)^T$ i $\mathring{F}(0) = \{x \in \mathbb{R}^2 \mid x_1 < 0\}$. Lako se vidi da je $x^* = 0$ točka lokalnog (čak i globalnog) minimuma funkcije f za bilo koji skup S iz Primjera 2.16-2.19 i zadovoljava nužan geometrijski uvjet optimalnosti $\mathring{F}(0) \cap T_S(0) = \emptyset$. U primjerima koji slijede u nastavku ćemo rješavati zadaću minimizacije upravo za tako definiranu funkciju cilja f.*

Primjer 2.30 (Nastavak primjera 2.16)**.** *Fritz Johnov sustav se u ovom primjeru svodi na:*

$$\begin{aligned} \mu_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & -2 \\ 0 & 0 \end{pmatrix} \mu &= 0, \\ (\mu_0, \mu^T)^T &\gneqq 0, \end{aligned}$$

pri čemu je vektor $\mu \in \mathbb{R}^2$ Lagrangeov multiplikator za funkcije ograničenja $g_1(x) := -x_1 \leq 0$ i $g_2(x) := (x_1 - 1)^2 + x_2^2 - 1 \leq 0$. Primjetimo da nije potrebno pisati uvjet komplementarnosti jer su u ovom slučaju oba uvjeta zadovoljena u vidu jednakosti u točki $x = 0$ pa je jednadžba (6b) automatski zadovoljena za sve μ. Rješenje sustava je par (μ_0, μ), gdje je $\mu = (\mu_1, \frac{\mu_0 - \mu_1}{2})^T$, za svaki $\mu_0 > 0$, $0 \leq \mu_1 \leq \mu_0$.

Primjer 2.31 (Nastavak primjera 2.17)**.** *Analogno kao u prethodnom primjeru, Fritz Johnov sustav u točki $x^* = 0$ se svodi na:*

$$\begin{aligned} \mu_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \mu &= 0, \\ (\mu_0, \mu^T)^T &\gneqq 0, \end{aligned}$$

pri čemu je $\mu \in \mathbb{R}^3$. Rješenje sustava je svaki par (μ_0, μ), gdje je $\mu = (\mu_0, 0, \mu_3)^T$ za svaki $\mu_0 \geq 0$, $\mu_3 \geq 0$, uz uvjet da je barem jedan strogo veći od nule.

Primjer 2.32 (Nastavak primjera 2.18)**.** *U ovom primjeru se Fritz Johnov sustav u točki $x^* = 0$ svodi na:*

$$\begin{aligned} \mu_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix} \mu &= 0, \\ (\mu_0, \mu^T)^T &\gneqq 0, \end{aligned}$$

pri čemu je $\mu \in \mathbb{R}^3$. Rješenje sustava je par $(0, \mu)$, gdje je $\mu = (\mu_1, \mu_2, \mu_1 - \mu_2)^T$, za svaki $\mu_1 > 0$, $0 \leq \mu_2 \leq \mu_1$.

Primjer 2.33 (Nastavak primjera 2.19). *Fritz Johnov sustav u točki $x^* = 0$ se svodi na:*

$$\begin{aligned} \mu_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & 0 \end{pmatrix} \mu &= 0, \\ (\mu_0, \mu^T)^T &\gneqq 0, \end{aligned}$$

pri čemu je $\mu \in \mathbb{R}^3$. Rješenje sustava je svaki par (μ_0, μ), gdje je $\mu = (0, \mu_2, \mu 2 - \frac{\mu_0}{2})^T$ za svaki $\mu_2 > 0$, $0 \leq \mu_0 \leq 2\mu_2$.

Glavni nedostatak Fritz Johnovih uvjeta je što su preslabi. Također, primijetimo da u Fritz Johnovom sustavu može vrijediti $\mu_0 = 0$, a to znači da uvjeti opisuju samo funkcije ograničenja te su neovisni od funkcije cilja f koju želimo minimizirati. Naravno, takve situacije su nepoželjne pa u ostatku poglavlja opisujemo kako ih izbjeći.

Budući da konus dopustivih smjerova $R_S(x)$ može biti loša aproksimacija tangencijalnog konusa $T_S(x)$, a po Propoziciji 2.21 isto vrijedi i za $\mathring{G}(x)$, u općenitom slučaju ne možemo poboljšati Fritz Johnove uvjete optimalnosti. Ipak, uz dodatnu pretpostavku regularnosti skupa S, to jest ako neki od skupova $\mathring{G}(x)$ ili $G(x)$ dovoljno dobro aproksimira skup $T_S(x)$, to možemo postići. Ovakve dodatne zahtjeve ćemo zvati uvjeti regularnosti i označavat ćemo ih s CQ.

Sljedeća propozicija govori da u slučaju kada je zadovoljen uvjet regularnosti $\mathring{G}(x^*) \neq \emptyset$ bilo koje rješenje Fritz Johnovog sustava mora zadovoljavati $\mu_0 \neq 0$.

Propozicija 2.34. *Neka je $x^* \in S$ točka lokalnog minimuma i neka vrijedi $\mathring{G}(x^*) \neq \emptyset$. Tada multiplikator μ_0 u Fritz Johnovom sustavu ne može biti jednak nuli.*

Dokaz. Tvrdnju dokazujemo obratom po kontrapoziciji. Pretpostavimo da za rješenje Fritz Johnovog sustava vrijedi $\mu_0 = 0$. Promotrimo matricu A definiranu na sljedeći način:

$$A = (\nabla g_i(x^*)), \quad i \in I(x^*).$$

Budući da je $A\mu = 0$, $\mu \geq 0$ i $\mu \neq 0$, po Farkasevoj lemi slijedi da sustav $A^T p < 0$ nema rješenja, to jest vrijedi $\mathring{G}(x^*) = \emptyset$.

∎

Napomena 2.35. *Dijeljenjem svih jednadžbi sustava* (6) *s $\mu_0 > 0$ dolazimo do pretpostavke $\mu_0 = 1$ i u tom slučaju će biti zadovoljeni Karush-Kuhn-Tuckerovi uvjeti koje proučavamo u nastavku.*

2.5 Karush–Kuhn–Tuckerovi uvjeti

U ovom odjeljku ćemo iskazati Karush-Kuhn-Tuckerove uvjete optimalnosti koji su specijalan slučaj Fritz Johnovih za $\mu_0 = 1$. Promatrat ćemo istu zadaću nelinearnog programiranja kao i do sada, a kasnije ćemo diskutirati o slučaju kada dodatno imamo uvjete tipa jednakosti. Kako bismo došli do KKT uvjeta uvodimo dodatni zahtjev regularnosti na dopustivi skup S. Za početak ćemo pretpostaviti da je zadovoljen Abadijev uvjet regularnosti, koji je u praksi ipak vrlo teško provjeriti. Kasnije u radu ćemo iskazati neke druge pretpostavke na dopustivi skup koje su lakše za provjeriti, a impliciraju ovu. Započnimo formalnom definicijom.

Definicija 2.36. *Kažemo da je u točki $x \in S$ zadovoljen* ***Abadijev uvjet regularnosti*** *(Abadi CQ) ako vrijedi $T_S(x) = G(x)$.*

Primjer 2.37. *Od primjera koje smo do sada proučavali samo Primjeri 2.16 i 2.19 zadovoljavaju Abadijev CQ.*

Spremni smo za dokaz najvažnijeg teorema ovog rada.

Teorem 2.38 (Karush-Kuhn-Tuckerovi[4] uvjeti optimalnosti)**.** *Neka je $x^* \in S$ točka lokalnog minimuma. Ako je u točki x^* zadovoljen Abadijev uvjet regularnosti onda postoji vektor $\mu \in \mathbb{R}^m$ takav da vrijedi*

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0, \tag{7a}$$

$$\mu_i g_i(x^*) = 0, \quad i = 1, \dots, m, \tag{7b}$$

$$\mu \geq 0. \tag{7c}$$

Uvjete (7a) – (7c) nazivamo Karush-Kuhn-Tuckerovi uvjeti optimalnosti, ili skraćeno KKT uvjeti. Točku koja ih zadovoljava zovemo KKT točka. Slično kao i kod Fritz Johnovih uvjeta, vektor μ koji zadovoljava jednadžbe se naziva Lagrangeov, ili ponekad, KKT multiplikator.

Dokaz. Zbog nužnog geometrijskog uvjeta optimalnosti vrijedi $\mathring{F}(x^*) \cap T_S(x^*) = \emptyset$, što uz pretpostavku Abadijevog uvjeta regularnosti u točki $x^* \in S$ implicira $\mathring{F}(x^*) \cap G(x^*) = \emptyset$. Kao i u dokazu teorema Fritz Johnovih uvjeta optimalnosti (Teorem 2.28) konstruiramo matricu A na sljedeći način:

$$A = \left(\nabla g_i(x^*)\right), \quad i \in I(x^*).$$

Tada sustav $A^T p \leq 0$, $-\nabla f(x^*)^T p > 0$ nema rješenja. Po Farkasevoj lemi slijedi da sustav $A\xi = -\nabla f(x^*)$, $\xi \geq 0$ ima rješenje. Definirajmo vektor μ na sljedeći način: $\mu_{I(x^*)} = \xi$ i $\mu_i = 0, i \notin I(x^*)$. Tako definirani vektor μ zadovoljava KKT uvjete (7).

■

[4]William Karush (1. ožujak 1917. – 22. veljače 1997.), američki matematičar i fizičar; Harold William Kuhn (29. srpanj 1925. – 2. srpanj 2014.), američki matematičar; Albert William Tucker (28. studeni 1905. – 25. siječanj 1995.), američki matematičar

Napomena 2.39. *Uočimo da se u zadaćama bezuvjetne optimizacije KKT sustav svodi samo na uvjet* $\nabla f(x^*) = 0$*, što smo već opisali Teoremom 2.1.*

Primjer 2.40 (Nastavak primjera 2.16)**.** *Ranije smo pokazali da u točki* $x^* = 0$ *vrijedi* $T_S(0) = G(0)$*, to jest zadovoljen je Abadijev uvjet regularnosti pa KKT sustav ima rješenje. Rješenje pripadnog sustava*

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & -2 \\ 0 & 0 \end{pmatrix} \mu &= 0, \\ \mu &\gneqq 0 \end{aligned}$$

je dano s $\mu = (\mu_1, \frac{1}{2}(1 - \mu_1))^T$*, za svaki* $0 \le \mu_1 \le 1$.

Primjer 2.41 (Nastavak primjera 2.17)**.** *U ovom primjeru nije zadovoljena pretpostavka Abadijevog uvjeta regularnosti, ali KKT sustav:*

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \mu &= 0, \\ \mu &\gneqq 0 \end{aligned}$$

ipak ima rješenje $\mu = (1, 0, \mu_3)^T$*, za svaki* $\mu_3 \ge 0$.

Primjer 2.42 (Nastavak primjera 2.18)**.** *Budući da u ovom primjeru za rješenje Fritz Johnovog sustava nužno vrijedi* $\mu_0 = 0$ *pripadni KKT sustav:*

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix} \mu &= 0, \\ \mu &\gneqq 0 \end{aligned}$$

nema rješenja. Nije zadovoljena ni pretpostavka Abadijevog uvjeta regularnosti.

Primjer 2.43 (Nastavak primjera 2.19)**.** *U ovom primjeru je zadovoljena pretpostavka Abadijevog uvjeta regularnosti i KKT sustav:*

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -2 & 2 \\ -1 & 0 & 0 \end{pmatrix} \mu &= 0, \\ \mu &\gneqq 0 \end{aligned}$$

ima rješenje $\mu = (0, \mu_2, \mu_2 - \frac{1}{2})$*, za svaki* $\mu_2 \ge \frac{1}{2}$*. U ovom slučaju skup Lagrangeovih multiplikatora nije ograničen i to proizlazi iz toga što smo uvjet tipa jednakosti rastavili na dva uvjeta tipa nejednakosti. U nastavku rada ćemo formulirati KKT sustav koji će zadržati originalnu reprezentaciju skupa* S *i na taj način ćemo smanjiti broj multiplikatora za uvjete tipa jednakosti samo na jednoga.*

Uvjeti tipa jednakosti

Do sada smo razmatrali samo zadaće nelinearnog programiranja s uvjetima tipa nejednakosti, a sada želimo vidjeti kako se dokazana teorija može primijeniti u slučaju zadaća u kojima se javljaju i uvjeti tipa jednakosti.

Dakle, promatrajmo dopustivi skup S koji je definiran uvjetima tipa jednakosti i nejednakosti na sljedeći način:

$$S := \{\, x \in R^n \mid g_i(x) \leq 0,\ i = 1, \ldots, m;\ h_j(x) = 0,\ j = 1, \ldots, k \,\}, \tag{8}$$

pri čemu su $g_i \colon \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ te $h_j \colon \mathbb{R}^n \to \mathbb{R}, j = 1, \ldots, k$ neprekidno diferencijabilne funkcije ograničenja.

Kako bismo sveli ovaj problem na problem koji smo analizirali ranije, skup S možemo zapisati preko uvjeta tipa nejednakosti pomoću neprekidno diferencijabilnih funkcija $\widetilde{g_i} \colon \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m + 2k$ definiranih s:

$$\widetilde{g_i} := \begin{cases} g_i, & i = 1, \ldots, m, \\ h_{i-m}, & i = m + 1, \ldots, m + k, \\ -h_{i-m-k}, & i = m + k + 1, \ldots, m + 2k. \end{cases} .$$

Tada je ekvivalentan zapis dopustivog skupa

$$S := \{\, x \in R^n \mid \widetilde{g_i}(x) \leq 0,\ i = 1, \ldots, m + 2k \,\}. \tag{9}$$

Neka je za novu reprezentaciju skupa S sa $\widetilde{G}(x)$ označen konus

$$\widetilde{G}(x) = \{\, p \in R^n \mid \nabla\widetilde{g_i}(x)^T p \leq 0,\ i \in I(x) \,\},$$

a staru oznaku $G(x)$ koristimo za konus definiran samo gradijentima aktivnih funkcija ograničenja tipa nejednakosti u reprezentaciji (8), to jest

$$G(x) = \{\, p \in R^n \mid \nabla g_i(x)^T p \leq 0,\ i \in I(x) \,\}.$$

Dodatno, uvodimo potprostor definiran samo gradijentima funkcija ograničenja tipa jednakosti:

$$H := \{\, p \in R^n \mid \nabla h_i(x)^T p = 0,\ i = 1, \ldots, k \,\}.$$

Budući da su sve funkcije ograničenja $\widetilde{g_i}, i = m + 1, \ldots, m + 2k$ nužno aktivne u svakom $x \in S$, vrijedi

$$\widetilde{G}(x) = G(x) \cap H(x) \tag{10}$$

pa Abadijev uvjet regularnosti za skup (8) možemo ekvivalentno zapisati kao

$$T_S(x) = G(x) \cap H(x).$$

Ako pretpostavimo da je prethodni uvjet zadovoljen, KKT sustav za reprezentaciju skupa S danu s (9) možemo zapisati u sljedećem obliku:

$$\begin{aligned}\sum_{i=1}^{m}\mu_i\nabla g_i(x^*)+\sum_{i=m+1}^{m+k}\mu_i\nabla h_{i-m}(x^*)-\sum_{i=m+k+1}^{m+2k}\mu_i\nabla h_{i-m-k}(x^*)&\\ +\nabla f(x^*)&=0, &&(11a)\\ \mu_i g_i(x^*)&=0,\quad i=1,\dots,m, &&(11b)\\ \mu_i h_{i-m}(x^*)&=0,\quad i=m+1,\dots,m+k, &&(11c)\\ -\mu_i h_{i-m-k}(x^*)&=0,\quad i=m+k+1,\dots,m+2k, &&(11d)\\ \mu&\geq 0. &&(11e)\end{aligned}$$

Definirajmo par vektora $(\widetilde{\mu},\widetilde{\lambda})\in\mathbb{R}^m\times\mathbb{R}^k$ sa $\widetilde{\mu}_i=\mu_i, i=1,\dots,m$ i $\widetilde{\lambda}_j=\mu_{m+j}-\mu_{m+k+j}, j=1,\dots,k$. Uočimo da uvjeti (11c) i (11d) nisu potrebni jer $x^*\in S$ implicira $h_j(x^*)=0, j=1,\dots,k$. Dakle, za zadaće u kojima je dopustivi skup opisan uvjetima tipa nejednakosti i jednakosti dobivamo sljedeći KKT sustav:

$$\nabla f(x^*)+\sum_{i=1}^{m}\widetilde{\mu}_i\nabla g_i(x^*)+\sum_{j=1}^{k}\widetilde{\lambda}_j\nabla h_j(x^*)=0, \tag{12a}$$

$$\widetilde{\mu}_i g_i(x^*)=0,\quad i=1,\dots,m, \tag{12b}$$

$$\widetilde{\mu}\geq 0. \tag{12c}$$

Ovime smo pokazali da vrijedi tvrdnja sljedećeg teorema.

Teorem 2.44 (KKT uvjeti optimalnosti za uvjete tipa jednakosti i nejednakosti)**.** *Pretpostavimo da je dopustivi skup S definiran s* (8) *i neka je $x^*\in S$ točka lokalnog minimuma. Ako je u točki x^* zadovoljen Abadijev uvjet regularnosti onda postoji par vektora $(\widetilde{\mu},\widetilde{\lambda})\in\mathbb{R}^m\times\mathbb{R}^k$ koji zadovoljava sustav* (12).

Napomena 2.45. *Primijetimo da u sustavu nema uvjeta na predznak multiplikatora $\widetilde{\lambda}$ koji se odnosi na uvjete tipa jednakosti.*

Primjer 2.46 (Nastavak primjera 2.43)**.** *U Primjeru 2.43 smo računali Lagrangeove multiplikatore nakon što smo dopustivi skup S zapisali samo pomoću uvjeta tipa nejednakosti. Izračunajmo ih sada za originalnu reprezentaciju dopustivog skupa, $S=\{x\in\mathbb{R}^2\mid x_2\geq 0;\ (x_1-1)^2+x_2^2=1\}$. Zadovoljen je Abadijev uvjet regularnosti i optimalna točka postoji pa KKT sustav nužno ima rješenje. Laganim računom se dobije da je rješenje sustava*

$$\begin{aligned}\begin{pmatrix}1\\0\end{pmatrix}+\mu_1\begin{pmatrix}0\\-1\end{pmatrix}+\lambda_1\begin{pmatrix}-2\\0\end{pmatrix}&=0,\\ \mu_1&\geqq 0\end{aligned}$$

jedinstveno i vrijedi $\mu_1=0,\lambda_1=\frac{1}{2}$.

Uvjeti regularnosti

Pokazali smo da su KKT uvjeti zadovoljeni ako je zadovoljen Abadijev uvjet regularnosti. Međutim, taj uvjet nije uvijek lako provjeriti. Stoga u ovom pododjeljku proučavamo druge uvjete regularnosti koji impliciraju Abadijev i praktičniji su za korištenje. Pretpostavit ćemo da je dopustivi skup S definiran s

$$S = \{ x \in R^n \mid g_i(x) \leq 0,\ i = 1, \ldots, m;\ h_j(x) = 0,\ j = 1, \ldots, k \}.$$

Definicija 2.47. *Kažemo da je u točki $x \in S$ zadovoljen* ***Mangasarian-Fromovitzov uvjet regularnosti*** *(MFCQ) ako su gradijenti $\nabla h_j(x)$, $j = 1, \ldots, k$ linearno nezavisni i vrijedi $\mathring{G}(x) \cap H(x) \neq \emptyset$.*

Propozicija 2.48. *Ako vrijedi MFCQ onda vrijedi Abadijev CQ.*

Dokaz. Treba pokazati sljedeće:

(i) $\mathrm{cl}\,(\,\mathring{G}(x) \cap H(x)\,) \subseteq T_S(x)$

(ii) $T_S(x) \subseteq G(x) \cap H(x) = \widetilde{G}(x)$

(iii) $\mathrm{cl}\,(\mathring{G}(x) \cap H(x)\,) = G(x) \cap H(x)$

Tada očito vrijedi $T_S(x) = G(x) \cap H(x)$ što je Abadijev uvjet regularnosti. Dokažimo potrebne tvrdnje:

(ii) Tvrdnja slijedi iz Propozicije 2.21.

(iii) $\boxed{\subseteq}$ $G(x) \cap H(x)$ je zatvoren i vrijedi $\mathring{G}(x) \cap H(x) \subseteq G(x) \cap H(x)$ pa tvrdnja slijedi iz definicije zatvarača.
$\boxed{\supseteq}$ Neka je $p \in G(x) \cap H(x)$ te $\widetilde{p} \in \mathring{G}(x) \cap H(x) \neq \emptyset$. Tada je $p_n := (1 - \frac{1}{n})p + \frac{1}{n}\widetilde{p} \in \mathring{G}(x) \cap H(x)$ te $p_n \to p$ kada $n \to +\infty$.

(i) Označimo s $M = \{ x \in \mathbb{R}^n \mid h_j(x) = 0, j = 1, \ldots, k \}$. Kako je $T_S(x)$ zatvoren, dovoljno je pokazati $\mathring{G}(x) \cap H(x) \subseteq T_S(x)$. Neka je $\widetilde{p} \in \mathring{G}(x) \cap H(x)$. Za točku y iz neke okoline točke x vrijedi da su gradijenti $\nabla h_j(y)$, $j = 1, \ldots, k$ linearno nezavisni. Označimo s $\Pi(y) \colon \mathbb{R}^n \to \mathbb{R}^n$ linearan operator ortogonalne projekcije na potprostor $H(y)$, gdje je

$$H(y) = [\,\{ \nabla h_1(y), \ldots, \nabla h_k(y) \}\,]^{\perp}.$$

Promotrimo sada sustav običnih diferencijalnih jednadžbi

$$\begin{cases} \frac{d\gamma}{dt}(t) = \Pi(\gamma(t))\widetilde{p} \\ \gamma(0) = x \end{cases}$$

Preslikavanje $y \mapsto \Pi(y)$ je neprekidno jer je h neprekidno diferencijabilna pa prema Peanovom teoremu postoji rješenje $\gamma\colon \langle -\epsilon, \epsilon\rangle \to \mathbb{R}^n$. Tvrdimo da za dovoljno mali $\bar{\epsilon} \leq \epsilon$ vrijedi

$$\gamma(t) \text{ je dopustiva točka za } t \in [0, \bar{\epsilon}\rangle.$$

Za uvjete tipa nejednakosti imamo:
Ako je $i \notin I(x)$ onda vrijedi $g_i(x) < 0$ te zbog neprekidnosti funkcija g_i i γ vrijedi $g_i(\gamma(t)) < 0$ za dovoljno mali $|t|$. Za $i \in I(x)$ vrijedi

$$\frac{d}{dt} g_i(\gamma(t)) = \nabla g_i(\gamma(t))^T \gamma'(t) = \nabla g_i(\gamma(t))^T \Pi(\gamma(t))\widetilde{p}.$$

Dakle, vrijedi

$$\frac{d}{dt} g_i(\gamma(t))\Big|_{t=0} = \nabla g_i(x)^T \underbrace{\Pi(x)\widetilde{p}}_{=\,\widetilde{p} \text{ jer } \widetilde{p} \in H(x)} = \nabla g_i(x)^T \widetilde{p} < 0,$$

gdje posljednja nejednakost vrijedi zbog $\widetilde{p} \in \mathring{G}(x)$. Dakle, $t \mapsto g_i(\gamma(t))$ strogo pada za dovoljno mali $|t|$ pa je $g_i(\gamma(t)) < 0$, $t \in [0, \bar{\epsilon}\rangle$.

Za uvjete tipa jednakosti imamo:
Prema Lagrangeovom teoremu srednje vrijednosti, za $s \in \langle 0, t\rangle$ imamo

$$\begin{aligned} h_j(\gamma(t)) &= \underbrace{h_j(\gamma(0))}_{=0} + t\frac{d}{dt} h_j(\gamma(s)) \\ &= t\nabla h_j(\gamma(s))^T \gamma'(s) \\ &= t\nabla h_j(\gamma(s))^T \underbrace{\Pi(\gamma(s))\widetilde{p}}_{\in H(\gamma(s))} \\ &= 0. \end{aligned}$$

Dakle, $\gamma(t) \in M$ za $t \in \langle -\epsilon, \epsilon\rangle$.
Sada u definiciji $T_S(x)$ uzmemo $x_k = \gamma\left(\frac{1}{k}\right)$ i $\lambda_k = k$ pa vrijedi

$$\begin{aligned} &\lim_{k\to+\infty} x_k = \gamma(0) = x \\ &\lim_{k\to+\infty} k\left(\gamma\left(\frac{1}{k}\right) - x\right) = \lim_{k\to+\infty} \frac{\gamma\left(\frac{1}{k}\right) - \gamma(0)}{\frac{1}{k}} = \gamma'(0) = \widetilde{p} \end{aligned}$$

iz čega slijedi $\widetilde{p} \in T_S(x)$.

■

Primjer 2.49. *Ranije smo pokazali da Primjeri 2.17 i 2.18 ne zadovoljavaju Abadijev CQ pa budući da MFCQ implicira Abadijev CQ, obratom po kontrapoziciji zaključujemo da je u istim primjerima nužno narušen i MFCQ. S druge strane, Primjeri 2.16 i 2.19 zadovoljavaju MFCQ (a time i Abadijev CQ).*

Definicija 2.50. *Kažemo da je zadovoljen* ***Slaterov uvjet regularnosti*** *(Slater CQ) ako su funkcije* $g_i, i = 1, \dots, m$ *konveksne,* h_j, $j = 1, \dots, k$ *afine te postoji* $\bar{x} \in S$ *takva da* $g_i(\bar{x}) < 0$, $i = 1, \dots, m$.

Propozicija 2.51. *Ako vrijedi Slaterov uvjet regularnosti onda vrijedi MFCQ.*

Dokaz. Pretpostavimo da je u točki $\bar{x} \in S$ zadovoljen Slaterov uvjet regularnosti. Iz konveksnosti funkcija $g_i, i = 1, \dots, m$ slijedi

$$0 > g_i(\bar{x}) = g_i(\bar{x}) - g_i(x) \geq \nabla g_i(x)^T (\bar{x} - x), \quad i \in I(x).$$

Budući da su uvjeti tipa jednakosti afine funkcije slijedi

$$0 = h_j(\bar{x}) - h_j(x) = \nabla h_j(x)^T (\bar{x} - x), \quad j = 1, \dots, k.$$

Definirajmo $d := \bar{x} - x$. Tada $d \in \mathring{G}(x) \cap H(x)$, to jest vrijedi $\mathring{G}(x) \cap H(x) \neq \emptyset$.

■

Primjer 2.52. *Samo je u Primjeru 2.16 zadovoljen Slaterov CQ. Funkcije* g_i, $i = 1, 2$ *su konveksne i primjerice jedna Slaterova točka je* $(1, 0) \in S$ *jer vrijedi* $g_1(1, 0) = -1 < 0$ *i* $g_2(1, 0) = (1 - 1)^2 + 0^2 - 1 = -1 < 0$.

Definicija 2.53. *Kažemo da je u točki* $x \in S$ *zadovoljen* ***uvjet linearne nezavisnosti*** *(LICQ) ako* $\nabla g_i(x), i \in I(x)$ *i* $\nabla h_j(x)$, $j = 1, \dots, k$ *čine linearno nezavisan skup.*

Propozicija 2.54. *Ako vrijedi LICQ onda vrijedi MFCQ.*

Dokaz. Neka je u točki $x \in S$ zadovoljen LICQ. Gradijenti $\nabla h_j(x)$, $j = 1, \dots, k$ su linearno nezavisni po pretpostavci pa ostaje pokazati da vrijedi $\mathring{G}(x) \cap H(x) \neq \emptyset$. Dakle, treba pronaći vektor $d \in \mathbb{R}^n$ takav da $\nabla g_i(x)^T d < 0$, $i \in I(x)$ i $\nabla h_j(x)^T d = 0$, $j = 1, \dots, k$. Neka su matrice G i H definirane na sljedeći način:

$$G = \Big(\nabla g_i(x)\Big),\ i \in I(x), \quad H = \Big(\nabla h_j(x)\Big),\ j \in 1, \dots, k.$$

Promotrimo matricu A definiranu s

$$A = \begin{pmatrix} G & H \end{pmatrix}^T \in M_{(|I(x)|+k) \times n}(\mathbb{R})$$

i vektor $b \in \mathbb{R}^{|I(x)|+k}$ definiran s $b = (-e, 0)^T$, gdje je $e \in \mathbb{R}^{|I(x)|}$ vektor jedinica. Kako su retci matrice A linearno nezavisni, sustav $Ad = b$ ima rješenje d koje zadovoljava MFCQ.

■

Ako su zadovoljeni LICQ, rješenje (μ, λ) Karush-Kuhn-Tuckerovog sustava (12) je nužno jedinstveno i upravo zbog toga je to najčešće korišten uvjet regularnosti.

Primjer 2.55. *Samo je u Primjeru 2.19 zadovoljen LICQ. Pokazali smo da za $x = (0, 0)$ vrijedi $\nabla g_1(0) = (0, -1)^T$ i $\nabla h_1(0) = (-2, 0)^T$. Očito su $\nabla g_1(0)$ i $\nabla h_1(0)$ linearno nezavisni. Također, primijetimo da smo za taj skup dobili jedinstveno rješenje Lagrangeovih multiplikatora (Primjer 2.46).*

Definicija 2.56. *Kažemo da je zadovoljen **uvjet afine regularnosti** ako su sve funkcije $g_i, i = 1, \dots, m$ i $h_j, j = 1, \dots, k$ afine.*

Propozicija 2.57. *Ako vrijedi uvjet afine regularnosti onda vrijedi Abadijev CQ.*

Dokaz. Pretpostavimo da su sve funkcije $g_i, i = 1, \dots, m$ i $h_j, j = 1, \dots, k$ afine. Tada je dopustivi skup S poliedarski skup pa je konus dopustivih smjerova $R_S(x)$ jednak $G(x) \cap H(x)$. Po Propoziciji 2.15 imamo $R_S(x) \subseteq T_S(x)$. Iz Propozicije 2.21 slijedi $T_S(x) \subseteq \widetilde{G}(x) = G(x) \cap H(x)$, što zajedno s prethodnim povlači $T_S(x) = G(x) \cap H(x)$, to jest zadovoljen je Abadijev uvjet regularnosti.

∎

Dovoljnost KKT uvjeta

Do sada smo pokazali da su KKT uvjeti optimalnosti nužni, što znači da se općenito na temelju njih može ustanoviti da neka točka nije rješenje problema nelinearnog progamiranja. Međutim, pitamo se mogu li KKT uvjeti biti dovoljni za optimalnost. Odnosno, ako je KKT sustav rješiv u x^* možemo li zaključiti da je x^* optimalna točka? Vidjeli smo da je u zadaćama bezuvjetne optimizacije takav zaključak omogućavalo svojstvo konveksnosti. Sada ćemo pokazati da je i u zadaćama uvjetne optimizacije potrebna slična pretpostavka.

Za početak definirajmo problem konveksnog programiranja (konveksne optimizacije). Radi se o problemu minimizacije konveksne funkcije na konveksnom skupu.

Definicija 2.58. *Promatramo sljedeću zadaću nelinearnog programiranja:*

$$\begin{aligned} min\ & f(x) \\ & x \in S, \end{aligned}$$

pri čemu je dopustivi skup definiran s

$$S = \{ x \in R^n \mid g_i(x) \le 0,\ i = 1, \dots, m;\ h_j(x) = 0,\ j = 1, \dots, k \}.$$

*Kažemo da se radi o **zadaći konveksnog programiranja** ako su funkcija cilja f i funkcije $g_i, i = 1, \dots, m$ konveksne te funkcije $h_j, j = 1, \dots, k$ afine.*

Teorem 2.59 (Dovoljnost KKT uvjeta za konveksne zadaće)**.** *Ako je $x^* \in S$ KKT točka zadaće konveksnog programiranja onda je x^* točka globalnog minimuma.*

Dokaz. Neka je $x \in S$ proizvoljan. Zbog konveksnosti funkcija $g_i, i = 1, \dots, m$ vrijedi

$$-\nabla g_i(x^*)^T(x - x^*) \geq g_i(x^*) - g_i(x) = -g_i(x) \geq 0, \quad i \in I(x^*), \tag{13}$$

a zbog pretpostavke da su funkcije $h_j, j = 1, \dots, k$ afine vrijedi

$$-\nabla h_j(x^*)^T(x - x^*) = h_j(x^*) - h_j(x) = 0, \quad j = 1, \dots, k. \tag{14}$$

Koristeći konveksnost funkcije cilja, jednadžbe (12a) i (12b), nenegativnost Lagrangeovih multiplikatora $\mu_i, i \in I(x^*)$ te relacije (13) i (14) dobivamo sljedeću nejednakost:

$$\begin{aligned} f(x) - f(x^*) &\geq \nabla f(x^*)^T(x - x^*) \\ &= -\sum_{i \in I(x^*)} \mu_i \nabla g_i(x^*)^T(x - x^*) - \sum_{j=1}^{k} \lambda_j \nabla h_j(x^*)^T(x - x^*) \\ &\geq 0. \end{aligned}$$

Kako je $x \in S$ bio proizvoljan, slijedi da je x^* točka globalnog minimuma.

■

Kao posljedica prethodnog teorema i nužnosti KKT uvjeta ističe se sljedeći rezultat.

Teorem 2.60. *Neka je za zadaću konveksnog programiranja zadovoljen Slaterov uvjet regularnosti. Tada je x^* točka globalnog minimuma ako i samo ako je x^* KKT točka.*

Sljedeći primjer pokazuje da u iskazu KKT teorema ne možemo izostaviti pretpostavku Slaterovog uvjeta regularnosti.

Primjer 2.61. *Promatramo sljedeći problem:*

$$\begin{aligned} &min\ x_1 \\ &\quad x_1^2 + x_2 \leq 0, \\ &\quad -x_2 \leq 0. \end{aligned}$$

Radi se o zadaći konveksnog programiranja i jedina dopustiva točka jest $0 \in \mathbb{R}^2$. Štoviše, to je točka globalnog minimuma. Obje funkcije ograničenja su aktivne u toj točki, to jest vrijedi $g_1(0) = 0$ i $g_2(0) = 0$ pa Slaterov uvjet regularnosti nije zadovoljen. Lako je provjeriti da KKT sustav

$$\begin{aligned} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \mu &= 0, \\ \mu &\geqq 0, \end{aligned}$$

nema rješenja, iz čega vidimo da čak ni u zadaćama konveksnog programiranja koje ne zadovoljavaju Slaterov uvjet regularnosti optimalna točka problema ne mora zadovoljavati KKT uvjete.

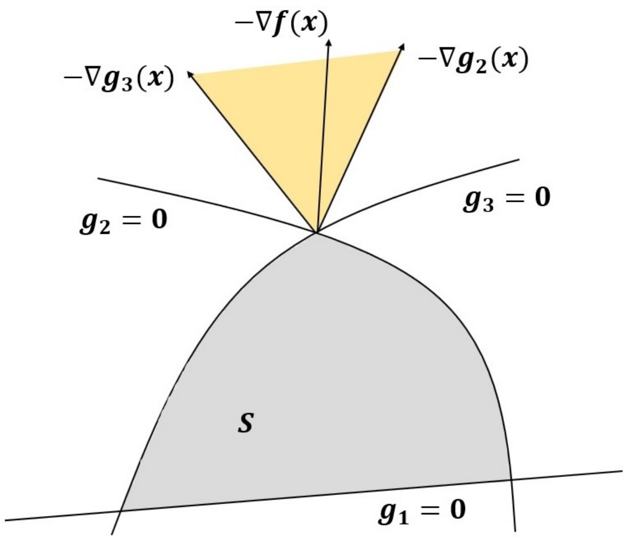
Napomena 2.62. *Primijetimo da u KKT sustavu uvjet komplementarnosti* $\mu_i g_i(x^*) = 0$ *povlači* $\mu_i = 0, i \notin I(x^*)$ *(za neaktivne funkcije ograničenja). Stoga KKT uvjete možemo pisati u sljedećem obliku:*

$$-\nabla f(\boldsymbol{x}^*) = \sum_{i \in I(x^*)} \mu_i \nabla g_i(\boldsymbol{x}^*),$$
$$\mu_i \geq 0, \quad i \in I(x^*).$$

Ovakva formulacija uvjeta govori da se negativni gradijent funkcije cilja u optimalnom rješenju može zapisati kao linearna kombinacija gradijenata aktivnih funkcija ograničenja u točki x^**, pri čemu su koeficijenti jednaki KKT mulitplikatoru. To znači da se negativni gradijent funkcije cilja u optimalnom rješenju nalazi u konusu s vrhom u točki* x^**, koji je u zadaćama konveksnog programiranja određen gradijentima aktivnih funkcija ograničenja u toj točki. Dakle, KKT sustav možemo geometrijski interpretirati kao*

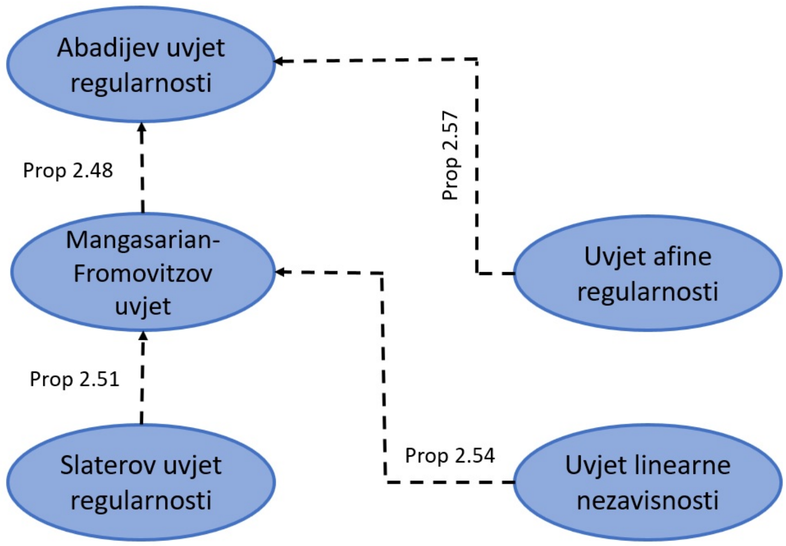
$$-\nabla f(x^*) \in N_S(x^*),$$

pri čemu je $N_S(x^*)$ *normalni konus za* S *u* $x^* \in S$ *(vidi Slika 2.6).*

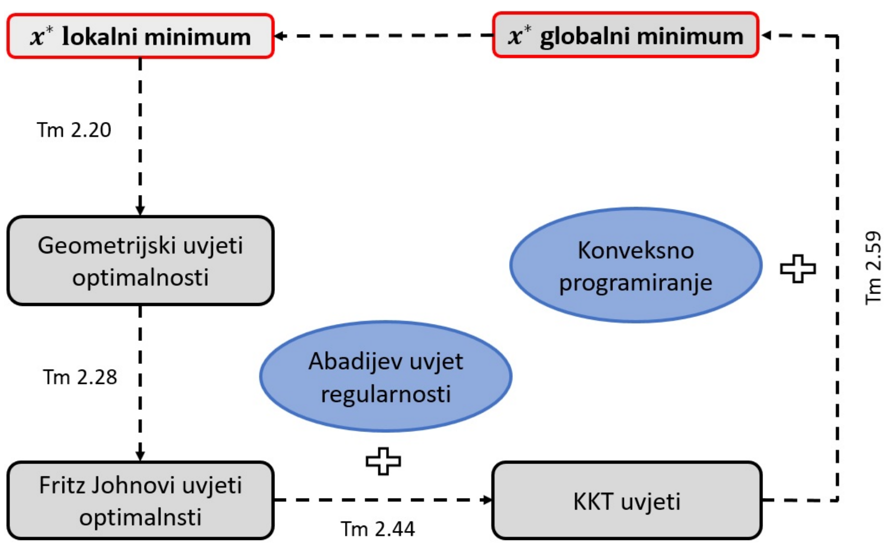


Slika 2.6: Grafička ilustracija KKT uvjeta

Na samom kraju ovog poglavlja, radi boljeg razumijevanja, na Slikama 2.7 i 2.8 je još jednom dan pregled najvažnijih dobivenih rezultata.



Slika 2.7: Veza između uvjeta regularnosti



Slika 2.8: Shematski prikaz poglavlja

Poglavlje 3

Primjena Karush-Kuhn-Tuckerovih uvjeta

Karush-Kuhn-Tuckerovi uvjeti optimalnosti primjenjuju se u širokom spektru područja. Koriste se pri rješavanju raznih matematičkih, ekonomskih ili financijskih problema. U ovom poglavlju ćemo na nekoliko primjera ilustrirati njihovu važnost. Započinjemo primjerom iz područja financija, a radi se o problemu izbora portfelja investicija.

Primjer 3.1. *Na financijskim tržištima većina pojedinaca i institucija ima u svome vlasništvu više različitih investicija, odnosno više različitih imovina, koje skupa čine portfelj. Upravljanje portfeljima vrijednosnica je jedno od temeljnih izazova svih financijskih institucija i cilj je pronaći način kako efikasno raspodijeliti svoja sredstva da bi se ostvario najveći mogući prinos uz što manji rizik. Svaki investitor je suprostavljen između dva cilja: visokog očekivanog prinosa i niskog rizika. U skladu s tim, model optimizacije portfelja se može formulirati na dva načina. Jedan način je da investitor odredi donju granicu prinosa portfelja i potom iz skupa mogućih portfelja izabere onaj koji ima minimalni rizik. Drugi način je da investitor unaprijed zada gornju granicu prihvatljivog rizika i onda iz skupa mogućih portfelja izabere onaj koji maksimizira prinos. U financijama je uobičajeno za mjeru rizika gledati drugi korijen iz varijance koji nazivamo standardnom devijacijom, a označavamo sa σ_P. Vrijednost standardne devijacije, kao i očekivanog srednjeg povrata dobiva se na osnovu povijesnih podataka.*

Pretpostavimo da investitor ulaže u dva tipa dionica: A i B. Maksimalan iznos novca kojeg može uložiti iznosi 1 000 *kuna. Neka je x_1 iznos koji investitor ulaže u dionicu A, a x_2 iznos koji ulaže u dionicu B. Neka je očekivani povrat ulaganja u portfelj dan sljedećom formulom:*

$$E[R_P] = 0.10x_1 + 0.12x_2,$$

dok je varijanca opisana s

$$\sigma_P^2 = 0.005x_1^2 - 0.006x_1x_2 + 0.008x_2^2.$$

Investitor želi raspodijeliti novac tako da zadrži standardnu devijaciju ne veću od 25, *ili ekvivalentno, varijancu ne veću od* 625. *Cilj je odrediti* x_1 *i* x_2 *kako bismo dobili portfelj koji maksimizira prinos, uz zadani uvjet gornje granice prihvatljivog rizika. Problem se može zapisati kao zadaća konveksnog programiranja*

$$\begin{aligned} \mathit{max}\ & 0.10x_1 + 0.12x_2 \\ & 0.005x_1^2 - 0.006x_1x_2 + 0.008x_2^2 - 625 \leq 0, \\ & x_1 + x_2 - 1000 \leq 0, \end{aligned}$$

ili ekvivalentno

$$\begin{aligned} \mathit{min}\ & -0.10x_1 - 0.12x_2 \\ & 0.005x_1^2 - 0.006x_1x_2 + 0.008x_2^2 - 625 \leq 0, \\ & x_1 + x_2 - 1000 \leq 0. \end{aligned}$$

Lako je provjeriti da su funkcije ograničenja i funkcija cilja zaista konveksne. Primjerice prva funkcija ograničenja se može zapisati u obliku $x^T Ax$, *pri čemu je* A *simetrična pozitivno definitna* 2×2 *matrica.*
Uočimo da je $(0, 0)$ *jedna Slaterova točka. Pripadni KKT sustav jednak je*

$$\begin{aligned} -0.10 + \mu_1(0.01x_1 - 0.006x_2) + \mu_2 &= 0, \\ 0.12 + \mu_1(-0.006x_1 + 0.016x_2) + \mu_2 &= 0, \\ \mu_1(0.005x_1^2 - 0.006x_1x_2 + 0.008x_2^2 - 625) &= 0, \\ \mu_2(x_1 + x_2 - 1000) &= 0, \\ \mu_1, \mu_2 &\geq 0. \end{aligned}$$

Kako se radi o zadaći konveksnog programiranja koja zadovoljava Slaterov uvjet regularnosti, optimalno rješenje gornjeg problema će biti upravo KKT točka istog problema. Rješavanjem sustava dobijemo da je optimalno rješenje $(x_1^*, x_2^*) = (348.01, 270.01)$. *Dakle, investitor treba* 348.01 *kuna uložiti u dionicu A i* 270.01 *kuna u dionicu B. Takvom raspodjelom novca će ostvariti prinos od* 67.20 *kuna, a varijanca će biti jednaka* 625.

Sljedeća dva primjera ilustrirat će kako se Karush-Kuhn-Tuckerovi uvjeti mogu koristiti kao instrument kvalitativne ekonomske analize.

Primjer 3.2 (Optimalno određivanje cijene)**.** *Mnoge tvrtke suočene su sa situacijom kada potražnja za određenim proizvodom ili uslugom varira o dobu dana, to jest u nekim trenucima kapaciteti tvrtke su iskorišteni u potpunosti, dok je u drugim trenucima potražnja* *izrazito slaba pa neki kapaciteti ostaju nedovoljno iskorišteni. U cilju maksimiziranja dobiti, tvrtke žele odrediti optimalnu cijenu proizvoda ili usluge, ovisno o razini potražnje kupaca. Samo neki od primjera takvih situacija su određivanje dnevne i noćne telefonske tarife ili povećanje cijene smještaja tijekom vrhunca turističke sezone.*

Označimo sa $x_1, x_2, \ldots, x_{24}$ *i* $p_1, p_2, \ldots, p_{24}$ *potraživanu količinu usluge svakog sata u danu te pripadnu cijenu, respektivno. Pretpostavimo da tijekom cijelog dana postoji potražnja za uslugom, to jest* $x_i > 0$, $i = 1, 2, \ldots, 24$. *Sa* y *označimo dostupnu količinu usluge (ponudu) po satu. Neka funkcija* $C(x_1, x_2, \ldots, x_{24})$ *označava dnevni ukupni operativni trošak i* $g(y)$ *dnevni trošak kapitala (kapaciteta). Pretpostavimo da su granični*[5] *operativni troškovi,* $\frac{\partial C}{\partial x_i}$ *te granični troškovi kapitala,* $\frac{\partial g}{\partial y}$, *pozitivni. Dodatno se pretpostavlja da na tržšištu vlada savršena konkurencija, to jest vrijedi* $\frac{\partial p_i}{\partial x_i} = 0$. *Tvrtke žele odrediti optimalnu cijenu usluge u svakom trenutku kako bi maksimizirale ukupnu dobit u danu, to jest radi se o maksimizaciji funkcije*

$$\pi = \sum_{i=1}^{24} p_i x_i - C(x_1, x_2, \ldots, x_{24}) - g(y),$$

uz dodatna ograničenja

$$\begin{aligned} x_i &\leq y, \quad i = 1, 2, \ldots, 24, \\ x_i &\geq 0, \quad i = 1, 2, \ldots, 24. \end{aligned}$$

KKT sustav za ovu zadaću je

$$-p_i + \frac{\partial C}{\partial x_i} + \mu_i = 0, \quad i = 1, 2, \ldots, 24, \tag{16a}$$

$$\frac{\partial g}{\partial y} - \sum_{i=1}^{24} \mu_i = 0, \tag{16b}$$

$$(x_i - y)\mu_i = 0, \quad i = 1, 2, \ldots, 24, \tag{16c}$$

$$\mu_i \geq 0, \quad i = 1, 2, \ldots, 24. \tag{16d}$$

U trenucima t *u kojima je potražnja slabija, po definiciji vrijedi* $y > x_t$ *pa uvjet komplementarnosti* (16c) *povlači* $\mu_t = 0$ *za svaki period smanjene potražnje. Tada iz uvjeta* (16a) *dobivamo*

$$p_t = \frac{\partial C}{\partial x_t}, \quad \text{za svaki trenutak smanjenje potražnje } t,$$

[5]Granični trošak predstavlja porast troška koji nastaje zbog proizvodnje jedne dodatne jedinice proizvoda.

to jest zaključujemo da je optimalna cijena u takvim trenucima jednaka graničnim operativnim troškovima (sirovine, rad i slično). Budući da postoje neiskorištene usluge, potražnju treba poticati što je moguće nižim cijenama, a da ne dođe do gubitka po prodajnoj jedinici.
S druge strane, u trenucima s u kojima dolazi do povećane potražnje za uslugama, kapacitet tvrki je u potpunosti iskorišten, to jest vrijedi $x_s = y$. Budući da smo pretpostavili $\frac{\partial g}{\partial y} > 0$, iz (16b) *slijedi*

$$\frac{\partial g}{\partial y} = \sum_s u_s > 0,$$

to jest u barem nekom trenutku povećane potražnje Lagrangeov multiplikator mora biti pozitivan. Iz (16a) *zaključujemo da mora vrijediti*

$$p_s = \frac{\partial C}{\partial x_s} + \mu_s, \quad \text{za svaki trenutak povećane potražnje } s.$$

Dakle, u ovakvim trenucima će cijena premašiti granični operativni trošak za dodatni iznos koji je jednak vrijednosti Lagrangeovog multiplikatora μ_s. Štoviše, iz (16b) *zaključujemo da će suma vrijednosti razlike cijena iznad graničnih operativnih troškova za sva razdoblja povećane potražnje biti jednaka graničnom kapitalnom trošku $\frac{\partial g}{\partial y}$.*

Primjer 3.3 (Maksimizacija prihoda uz ograničenje dobiti)**.** *Pretpostavimo da tvrtka nudi samo jedan proizvod ili uslugu koji ćemo označiti s q i da je njegova prodaja pod utjecajem troškova oglašavanja koje ćemo označiti s a. Tvrtka želi maksimizirati svoj ukupni prihod $R(q, a)$ uz ograničenje dobiti*[6]*,*

$$\pi = R(q, a) - C(q) - a \geq m,$$

pri čemu je $C(q)$ trošak proizvodnje, a m minimalna dobit koju žele ostvariti. Pretpostavimo da su granični prihod od oglašavanja $\frac{\partial R}{\partial a}$ i granični trošak proizvodnje $\frac{\partial C}{\partial q}$, pozitivni. Pokažimo da je iznos koji maksimizira prihod takav da je prihod jednak zadanoj vrijednosti m, granični prihod $\frac{\partial R}{\partial q}$ pozitivan, a granična dobit $\frac{\partial \pi}{\partial q}$ negativan.

Radi se o sljedećem problemu

$$\begin{aligned} \max\, & R(q, a) \\ & R(q, a) - C(q) - a \geq m, \\ & q \geq 0, \\ & a \geq 0. \end{aligned}$$

[6]Dobit (ili profit) označava razliku između prihoda i troškova.

KKT sustav je

$$-\frac{\partial R}{\partial q} + \mu\left[-\frac{\partial R}{\partial q} + \frac{\partial C}{\partial q}\right] = 0, \tag{17a}$$

$$-\frac{\partial R}{\partial a} + \mu\left[-\frac{\partial R}{\partial a} + 1\right] = 0, \tag{17b}$$

$$\mu\big[- R(q,a) + C(q) + a + m\big] = 0, \tag{17c}$$

$$\mu \geq 0. \tag{17d}$$

Uz pretpostavku $q > 0$, iz jednadžbe (17a) *slijedi*

$$-\frac{\partial R}{\partial q}(1+\mu) = -\frac{\partial C}{\partial q}\mu,$$

to jest

$$\frac{\frac{\partial R}{\partial q}}{\frac{\partial C}{\partial q}} = \frac{\mu}{1+\mu}. \tag{18}$$

Jednadžbu (17b) *možemo ekvivalentno zapisati kao*

$$\frac{\partial R}{\partial a} = \frac{\mu}{1+\mu},$$

što uz pretpostavku $\frac{\partial R}{\partial a}$ povlači $\mu > 0$. Tada uvjet komplementarnosti implicira $-R(q,a) + C(q) + a + m = 0$, to jest vrijedi $\pi = m$. Kako je $0 < \frac{\mu}{1+\mu} < 1$, zbog pretpostavke $\frac{\partial C}{\partial q} > 0$, iz jednadžbe (18) *zaključujemo da mora vrijediti $\frac{\partial R}{\partial q} > 0$ i $\frac{\partial R}{\partial q} < \frac{\partial C}{\partial q}$. Konačno, iz prethodno pokazanoga slijedi da je granična dobit $\frac{\partial \pi}{\partial q} = \frac{\partial R}{\partial q} - \frac{\partial C}{\partial q}$, negativna.*

Bibliografija

[1] N. Andreasson, A. Evgrafov, M. Patriksson, *An introduction to continuous optimization: Foundations and fundamental algorithms*, Studentlitteratur, 2005.

[2] M. S. Bazaraa, H. D. Sherali, C. M. Shetty, *Nonlinear programming: Theory and algorithms*, Wiley-Interscience, 2006.

[3] J. Brinkhuis, V. Tikhomirov, *Optimization: Insights and applications*, Princeton University Press, 2005.

[4] L. Čaklović, *Geometrija linearnog programiranja*, Element, Zagreb, 2010.

[5] G. Giorgi, A. Guerraggio, J. Thierfelder, *Mathematics of optimization*, Elsevier, 2004.

[6] I. Gogić, P. Pandžić, J. Tambača, *Diferencijalni račun funkcija više varijabli*, dostupno na `https://web.math.pmf.unizg.hr/nastava/difraf/dif/Diferencijalni_racun.pdf` (kolovoz 2019.)

[7] M. Luptačik, *Mathematical optimization and economic analysis*, Springer, 2010.

[8] L. Neralić, *Uvod u matematičko programiranje 1*, Element, Zagreb, 2003.

[9] W. Sun, Y.-X. Yuan, *Optimization theory and methods: Nonlinear programming*, Springer, 2006.

[10] Š. Ungar, *Matematička analiza u* $\mathbb{R}^n$, Tehnička knjiga, Zagreb, 2005.

Sažetak

U ovom radu smo se bavili razvojem nužnih i dovoljnih uvjeta optimalnosti. Na samom početku smo obradili pitanje egzistencije rješenja zadaće nelinearnog programiranja i u tu svrhu smo dokazali poopćenu verziju Weierstrassovog teorema. Prokomentirali smo osnovne rezultate za jednostavniji slučaj uvjetne optimizacije u kojem je dopustivi skup zatvoren i konveksan. U ostatku rada analiziramo općenitiju zadaću nelinearnog programiranja u kojoj je dopustivi skup zatvoren i funkcija cilja diferencijabilna. Iskazali smo geometrijske uvjete optimalnosti koji zahtijevaju da je u točki lokalnog minimuma presjek tangencijalnog konusa i konusa svih smjerova silaska prazan skup. Budući da je tangencijalni konus općenito teško izračunati, uveli smo dodatne uvjete tipa nejednakosti na dopustivi skup koji olakšavaju računanje konusa koji dobro aproksimiraju tangencijalni konus. Novouvedeni konusi nas vode do Fritz Johnovih uvjeta optimalnosti koji su ipak dosta slabi i mogu ih zadovoljavati mnoge točke koje nisu lokalno optimalne. Stoga smo uveli dodatni zahtjev Abadijevog uvjeta regularnosti i time smo došli do specijalnog slučaja Fritz Johnovih uvjeta koji je poznat kao Karush-Kuhn-Tuckerovi uvjeti. Kako Abadijev uvjet nije uvijek lako provjeriti definirali smo i ostale najčešće korištene uvjete regularnosti koji impliciraju Abadijev. Pokazali smo da su Karush-Kuhn-Tuckerovi uvjeti optimalnosti općenito nužni, a u zadaćama konveksnog programiranja dovoljni za lokalnu optimalnost točke.

Summary

In this thesis we studied development of necessary and sufficient optimality conditions. At the beginning we introduced an important result about existence of solutions to a nonlinear programming problem, known as generalized Weierstrass' theorem. We introduced some basic results for the case of constrained optimization over closed and convex sets. The rest of the thesis deals with a quite general nonlinear programming problem, where feasible set is closed and the objective function is differentiable. We established geometric optimality conditions which say that at every point of local minimum, the intersection of tangent cone and cone of descent directions must be an empty set. Tangent cone is nearly impossible to compute for general feasible sets so we gave a specific description of a feasible set in terms of inequalities, which helps us to compute other cones that approximate tangent cone in many practical situations. In this way we obtain the Fritz John conditions that, however, are somewhat too weak to be practical and they can be satisfied by many points that have nothing in common with locally optimal points. We assume an additional regularity on feasible set, Abadie's constraint qualification, and we get special case of Fritz John conditions, known as Karush-Kuhn-Tucker optimality conditions. Abadie's constraint qualification is difficult to check when it comes to practical problems so we defined some computationally verifiable assumptions that all imply Abadie's constraint qualification. We showed that Karush-Kuhn-Tucker optimality conditions are necessary, and for convex problems the Karush-Kuhn-Tucker optimality conditions are sufficient for local optimality.

Životopis

Rođena sam 17. studenoga 1995. godine u Zagrebu. Osnovnu školu sam pohađala u Topuskom, a nakon toga srednju školu u Glini, smjer opća gimnazija. Nakon završenog srednjoškolskog obrazovanja, 2014. godine upisujem preddiplomski studij Matematika na Matematičkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu. Godine 2017. stekla sam diplomu sveučilišnog prvostupnika matematike i iste sam godine upisala diplomski sveučilišni studij Financijska i poslovna matematika na istom fakultetu.