FELTÉTELES OPTIMALIZÁLÁS ALGORITMUSAI

DR. NAGY TAMÁS
egyetemi docens

Miskolci Egyetem
Alkalmazott Matematikai Tanszék

„A bemutatott kutató munka a TÁMOP-4.2.1.B-10/2/KONV-2010-0001 jelű projekt részeként az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg”
„This research was carried out as part of the TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.”

Miskolc, 2012

Tartalomjegyzék

1. Bevezetés

Ebben a tananyagban a feltételes optimalizálás megoldási módszereivel foglalkozunk, kétféle módszercsaláddal ismerkedünk meg.

Az egyik módszercsalád alapgondolata az, hogy a feltételes optimalizálási feladat megoldását visszavezetjük feltétel nélküli optimalizálási feladatok sorozatos megoldására. Ezt a megoldási eljárást szokás SUMT (Sequential Unconstrained Minimization Technique) módszer néven is emlegetni. Két változatát ismertetjük, az egyik a büntetőfüggvényes eljárás, a másik a korlátfüggvényes eljárás. Mindkét módszer lényege az, hogy a feltételeket beépíti a célfüggvénybe, így kapunk egy feltétel nélküli optimalizálási feladatot, amelyet sorozatosan más-más paraméterrel oldunk meg.

A másik módszercsalád pedig a lehetséges irányok módszere. Ezeknél a módszereknél a feltétel nélküli optimalizálásban megismert vonalmenti minimalizálási módszert alkalmazzuk, lépésenként egy-egy irányban keressük a feladat célfüggvényének optimális megoldását. Két dologra kell ügyelnünk: Egyrészt olyan irányt kell alkalmaznunk, amely lehetséges (megengedett) irány, azaz legalább egy kicsit elmozdulva, a tartományban maradunk. Másrészt pedig az egyváltozós optimalizási feladatnál figyelembe kell venni, hogy az optimális megoldást úgy keressük, hogy a tartományban maradjunk, ez azt jelenti, hogy az egyváltozós optimalizási feladat is feltételes feladat lesz.

2. SUMT módszer

2.1. Büntetőfüggvényes eljárás

A büntetőfüggvényes módszer jobb megértése végett először tekintsünk egy olyan optimalizálási feladatot, ahol egyetlen egyenlőséges feltétel van:

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \text{min!} \\ h(\mathbf{x}) &= 0 \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

Most helyette tekintsük az alábbi feltétel nélküli optimalizálási feladatot:

$$\begin{aligned} f(\mathbf{x}) + \mu\,|h(\mathbf{x})| &\rightarrow \text{min!} \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

ahol $\mu > 0$ nagy valós szám. Ha μ elegendően nagy, akkor úgy érezzük, hogy olyan optimális megoldást kapunk, amelynél a $h(\mathbf{x}) \rightarrow 0$, azaz a feltétel megközelítőleg ki lesz elégítve. A büntetőfüggyvény elnevezés onnan származik, hogy az új célfüggvényben olyan tagot használunk, amely "büntet", ha a feltétel nem teljesedik.

Hasonló gondolattal kezelhetjük az olyan optimalizálási feladatot, amelyben egyetlen egyenlőtlenséges feltétel van:

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \text{min!} \\ g(\mathbf{x}) &\leqq 0 \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

Könnyen meggyőződhetünk arról, hogy itt a $\mu\,|g(\mathbf{x})|$ tag akkor is büntet, ha a feltétel teljesedi, azaz $g(\mathbf{x}) < 0$ esetben is. Itt többek között az alábbi feltétel nélküli optimalizálási feladatot javasolhatjuk:

$$\begin{aligned} f(\mathbf{x}) + \mu \max\{0, g(\mathbf{x})\} &\to \min! \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

A max{...} függvény olyan, hogy a két argumentuma közül a nagyobbikat kapja értékül, így valóban csak akkor van büntetés, ha a feltétel nem teljesül, azaz csak a $g(\mathbf{x}) > 0$ esetben.

Néhány jelölést bevezetünk. Az új célfüggvényt segédfüggvénynek (jelben ϕ), a $\mu > 0$ számot büntető paraméternek, a feltételekből képzett részt pedig büntetőfüggvénynek (jelben B), a μB mennyiséget pedig büntetőtagnak nevezzük. A segédfüggvény és az új feltétel nélküli optimalizálási feladat alakja a következő:

$$\begin{aligned} \phi(\mathbf{x}) &= f(\mathbf{x}) + \mu B(g(\mathbf{x}), h(\mathbf{x})) \to \min! \\ \mathbf{x} &\in \mathbb{R}^n \end{aligned}$$

A $B(y)$ büntetőfüggvénynek általánosságban olyan folytonos függvénynek kell lenni, hogy pozitív büntető értéket adjon, ha a feltételek nem teljesednek, amennyiben viszont a feltételek teljesednek, akkor zérus legyen a büntetés.

Egyenlőséges ($=$) feltétel esetén

$$B(y) = \begin{cases} = 0, & \text{ha } y = 0 \\ > 0, & \text{ha } y \neq 0 \end{cases}$$

Egyenlőtlenséges ($\leqq$) feltétel esetén

$$B(y) = \begin{cases} = 0, & \text{ha } y \leqq 0 \\ > 0, & \text{ha } y > 0 \end{cases}$$

Az előzőekben bemutattunk ilyen függvényeket, természetesen több ilyen tulajdonsággal rendelkező függvény is van, például egyenlőséges esetben y^2, egyenlőtlenséges esetben $y + |y|$ vagy $[\max\{0, y\}]^2$. Célszerű olyan büntetőfüggvényeket használni, amelyeknél a ϕ segédfüggvény differenciálható, így a feltétel nélküli optimalizálási feladat megoldására alkalmazhatjuk azokat a módszereket is, amelyek használják a gradiens ill. a Hesse mátrixot. Ezért a leggyakrabban az egyenlőséges esetben y^2, egyenlőtlenséges esetben pedig a $[\max\{0, y\}]^2$ büntetőfüggvényt használjuk.

Általánosságban, amikor több egyenlőtlenséges és egyenlőséges feltétel is van, azaz az optimalizálási feladat

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \quad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \quad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

alakú, akkor az alábbi formájú segédfüggvényt szokás használni

$$\phi(\mathbf{x}) = f(\mathbf{x}) + \mu \left[\sum_{i=1}^{m} [\max\{0, g_i(\mathbf{x})\}]^2 + \sum_{i=1}^{k} [h_i(\mathbf{x})]^2 \right].$$

A büntetőfüggvényes módszer lényege tehát az, hogy minden lépésben megoldjuk a $\phi(\mathbf{x}) \rightarrow \min!$ feltétel nélküli optimalizálási feladatot növekvő μ büntető paraméter értékekkel.

Az alábbiakban tekintsünk két mintapéldát.

1. Példa

Oldjuk meg az alábbi nagyon egyszerű optimalizálási feladatot büntetőfüggvényes módszerrel:

$$\begin{aligned} x &\rightarrow \min! \\ 1-x &= 0 \end{aligned}$$

A feladat egyszerűsége miatt azonnal látható, hogy az optimális megoldás $\bar{x}=1$, az optimális célfüggvényérték pedig $f(\bar{x})=1$.

A feladathoz tartozó segédfüggyény, aminek a minimumát keressük feltétel nélkül, az alábbi:

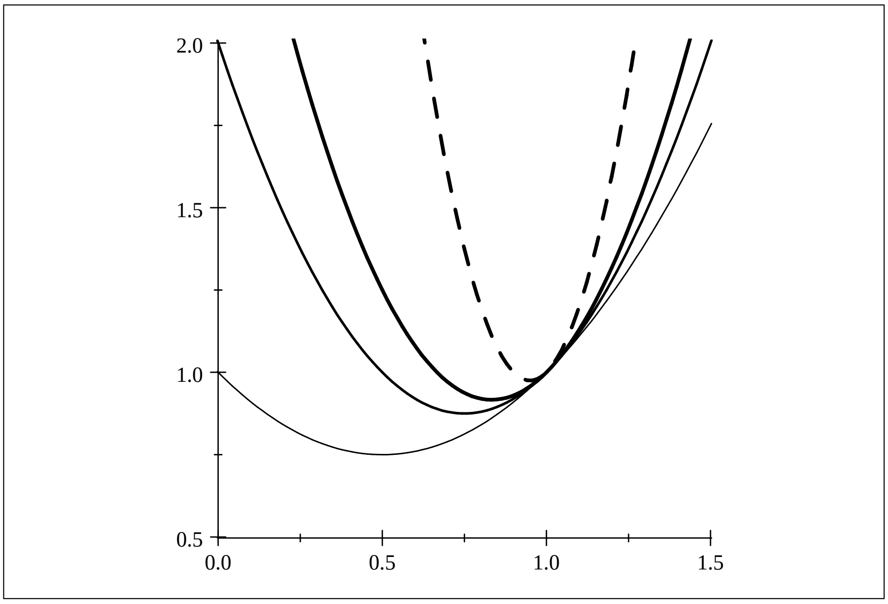
$$\phi(x)=f(x)+\mu(1-x)^2.$$

Ennek optimális megoldása:

$$x_\mu = 1-\frac{1}{2\mu}, \quad \mu>0$$

Könnyen ellenőrizhető, hogy ha $\mu \rightarrow +\infty$, akkor az $x_\mu \rightarrow 1$, a feltételes optimalizálási feladat optimumához, $\bar{x}$-höz.

Az alábbi ábra segítségével szemléltetjük a büntetőfüggvényes módszert. Az ábra a $\phi(x)$ segédfüggvényt ábrázolja $\mu=1, \mu=2, \mu=3, \mu=10$ büntetőparaméterek esetén. A legvékonyabban rajzolt függvény a $\mu=1$, a közepes vastagságú a $\mu=2$, a vastag a $\mu=3$, a szaggatottan rajzolt függvény pedig a $\mu=10$ büntetőparaméterhez tartozik.



Az ábrából nyomon követhetjük az alábbiakat: minél nagyobb a μ büntetőparaméter értéke, annál közelebb adódik a $\phi(x)$ segédfüggvény minimuma az eredeti feladat $f(x)$ célfüggvényének minimumához. Az ábra alapján az alábbi észrevételeket is tehetjük:

Ha $\mu \to +\infty$, akkor $x_\mu \to \bar{x} = 1$.
Ha $\mu \to +\infty$, akkor $f(x_\mu) \to f(\bar{x} = 1)$, monoton növekvő módon.
Ha $\mu \to +\infty$, akkor $\phi(x_\mu) \to f(\bar{x} = 1)$, monoton növekvő módon.
Ha $\mu \to +\infty$, akkor $\mu B(x_\mu) \to 0$, monoton csökkenő módon.
Ha $\mu \to +\infty$, akkor $B(x_\mu) \to 0$, monoton csökkenő módon.

2. Példa

Oldjuk meg az alábbi nagyon egyszerű optimalizálási feladatot büntetőfüggvényes módszerrel

$$\begin{aligned} x &\to \min! \\ 1 - x &\leqq 0 \end{aligned}$$

A feladat egyszerűsége miatt azonnal látható, hogy az optimális megoldás $\bar{x} = 1$, az optimális célfüggvényérték pedig $f(\bar{x}) = 1$.

A feladathoz tartozó segédfüggvény, aminek a minimumát keressük feltétel nélkül, az alábbi:

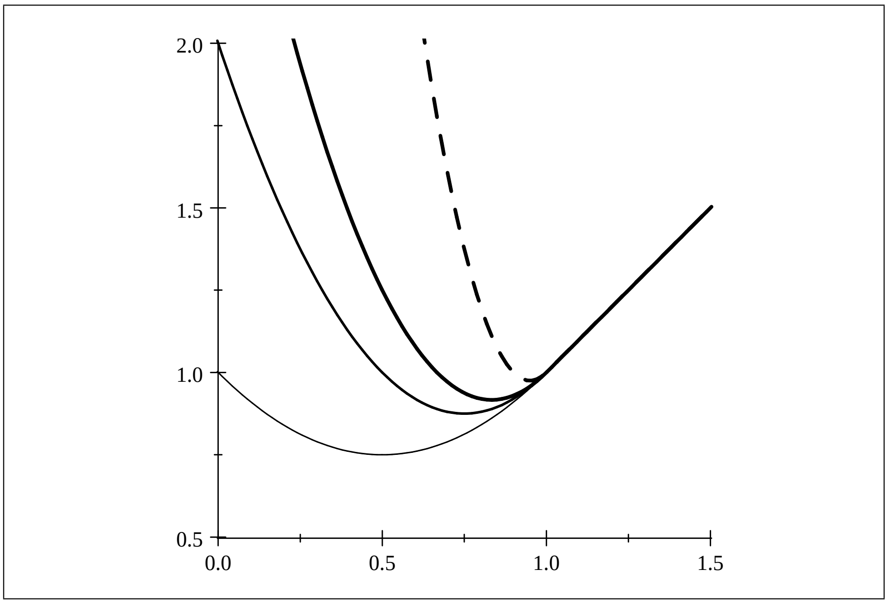
$$\phi(x) = f(x) + \mu\left[\max\{0, 1 - x\}\right]^2 = \begin{cases} x & \text{ha } x > 1 \\ x + \mu(1-x)^2 & \text{ha } x \leqq 1 \end{cases}$$

Ennek optimális megoldása:

$$x_\mu = 1 - \frac{1}{2\mu}, \quad \mu > 0$$

Könnyen ellenőrizhető, hogy ha $\mu \to +\infty$, akkor az $x_\mu \to 1$, a feltételes optimalizálási feladat optimumához, $\bar{x}$-höz.

Az alábbi ábra segítségével szemléltetjük a büntetőfüggvényes módszert. Az ábra a $\phi(x)$ segédfüggvényt ábrázolja $\mu = 1, \mu = 2, \mu = 3, \mu = 10$ büntetőparaméterek esetén. A legvékonyabban rajzolt függvény a $\mu = 1$, a közepes vastagságú a $\mu = 2$, a vastag a $\mu = 3$, a szaggatottan rajzolt függvény pedig a $\mu = 10$ büntetőparaméterhez tartozik.



Az ábrából nyomon követhetjük az alábbiakat: minél nagyobb a μ büntetőparaméter értéke, annál közelebb adódik a $\phi(x)$ segédfüggvény minimuma az eredeti feladat $f(x)$ célfüggvényének minimumához. Az előző példában tapasztalt észrevételeket ezen ábra alapján is megtehetjük, amelyek az alábbiak:

Ha $\mu \to +\infty$, akkor $x_\mu \to \bar{x} = 1$.
Ha $\mu \to +\infty$, akkor $f(x_\mu) \to f(\bar{x} = 1)$, monoton növekvő módon.
Ha $\mu \to +\infty$, akkor $\phi(x_\mu) \to f(\bar{x} = 1)$, monoton növekvő módon.
Ha $\mu \to +\infty$, akkor $\mu B(x_\mu) \to 0$, monoton csökkenő módon.
Ha $\mu \to +\infty$, akkor $B(x_\mu) \to 0$, monoton csökkenő módon.

A következő tétel a fenti észrevételek pontosítását adja:

TÉTEL
Tekintsük az alábbi optimalizálási feladat

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \quad i = 1, 2, \ldots, m \\ h_i(\mathbf{x}) &= 0 \quad i = 1, 2, \ldots, k \\ \mathbf{x} &\in X \end{aligned}$$

ahol f, g_i $i = 1, 2, \ldots, m;, h_i$ $i = 1, 2, \ldots, k$ folytonos függvények $\mathbb{R}^n$-ben, $X \subset \mathbb{R}^n$ nemüres halmaz. Tegyük fel, hogy a feladatnak van lehetséges megoldása és a $B(\mathbf{x})$ büntetőfüggvény folytonos függvény. Továbbá tegyük fel, hogy minden nemnegatív μ skalárra létezik $\mathbf{x}_\mu$ optimális megoldása a

$$\min\{\phi(\mathbf{x}) = f(\mathbf{x}) + \mu B(\mathbf{x}) : \mathbf{x} \in X\}$$

optimalizálási feladatnak és az $\{\mathbf{x}_\mu\}$ sorozat az X halmaz egy kompakt részhalmazában van. Ekkor

$$\begin{aligned} &\inf\{f(\mathbf{x}) : g_i(\mathbf{x}) \leqq 0, (i = 1, 2, \ldots, m), h_i(\mathbf{x}) = 0, (i = 1, 2, \ldots, k), \mathbf{x} \in X\} = \\ = \ &\sup\{\Theta(\mu) : \mu \geqq 0\} = \lim_{\mu \to \infty} \Theta(\mu), \end{aligned}$$

ahol $\Theta(\mu) = \inf\{\phi(\mathbf{x}) = f(\mathbf{x}) + \mu B(\mathbf{x}) : \mathbf{x} \in X\} = f(\mathbf{x}_\mu) + \mu B(\mathbf{x}_\mu)$. Továbbá az $\{\mathbf{x}_\mu\}$ sorozat bármely konvergens részsorozata tart az $\bar{\mathbf{x}}$ ponthoz, amely az eredeti optimalizálási feladat optimális megoldása. Ezenfelül, ha $\mu \to \infty$, akkor igazak az alábbiak:
a) $f(\mathbf{x}_\mu) \to f(\bar{\mathbf{x}})$ monoton növekvően,
b) $\phi(\mathbf{x}_\mu) \to f(\bar{\mathbf{x}})$ monoton növekvően,
c) $\Theta(\mu) \to f(\bar{\mathbf{x}})$ monoton növekvően,
d) $\mu B(\mathbf{x}_\mu) \to 0$ monoton csökkenően,
e) $B(\mathbf{x}_\mu) \to 0$ monoton csökkenően.

Az alábbiakban a büntetőfüggvényes módszer megoldási módját írjuk le.

Büntetőfüggvényes eljárás algoritmusa:
Induló lépés ($k = 1$): Kiindulunk egy tetszőleges $\mathbf{x}_k \in \mathbb{R}^n$ vektorból, választunk egy $\mu_k > 0$ büntetőparamétert és egy $\beta > 1$ skalárt.
Közbülső lépés:
1. Megoldjuk a $\phi(\mathbf{x}) = \min\{f(\mathbf{x}) + \mu_k B(\mathbf{x}) : \mathbf{x} \in X\}$ minimalizálási feladatot, legyen az optimális megoldás $\mathbf{x}_{k+1}$.

2. Megállunk, ha $\mu_k B(\mathbf{x}_{k+1}) < \varepsilon$, egyébként folytatjuk az eljárást a következő sorral.
3. Meghatározzuk a büntetőparaméter következő értékét, legyen $\mu_{k+1} = \beta\mu_k$.
4. Folytatjuk az eljárást, $k := k + 1$.

A gyakorlatban általában a $\mu_1 = 1$ büntetőparaméter értékkel indulunk és lépésenként a 10-szeresét vesszük, azaz a $\beta = 10$ értéket szokás választani, ezt tettük a következő példában is.

3. Példa
Oldjuk meg az alábbi optimalizálási feladatot büntetőfüggvényes módszerrel:

$$\begin{aligned} f(x_1, x_2) &= (x_1 - 1)^2 + x_2^2 \to \min! \\ h(x_1, x_2) &= x_1 - x_2 = 0 \end{aligned}$$

Megoldás
A feltétel nélküli optimalizálási segédfeladat:

$$\phi(x) = (x_1 - 1)^2 + x_2^2 + \mu \cdot (x_1 - x_2)^2 \to \min!$$

A kiinduló vektor legyen $\mathbf{x} = (10, 20)$, a kiinduló büntető paraméter pedig legyen $\mu = 1$. A megoldás egyes lépéseit az alábbi táblázatban közöljük:

μ	$\mathbf{x}_\mu$	$\phi(\mathbf{x}_\mu)$	$f(\mathbf{x}_\mu)$	$\mu B(\mathbf{x}_\mu)$	$B(\mathbf{x}_\mu)$
1	$\begin{bmatrix} 0.666 \\ 0.333 \end{bmatrix}$	0.333	0.222	0.111	0.111
10	$\begin{bmatrix} 0.524 \\ 0.476 \end{bmatrix}$	0.4762	0.4535	0.0227	0.00227
100	$\begin{bmatrix} 0.5025 \\ 0.4975 \end{bmatrix}$	0.4975	0.4950	0.002475	$2.475 \cdot 10^{-5}$
1000	$\begin{bmatrix} 0.500249 \\ 0.499749 \end{bmatrix}$	0.49975	0.49950	0.00025	$2.5 \cdot 10^{-7}$
10000	$\begin{bmatrix} 0.500024 \\ 0.4999974 \end{bmatrix}$	0.499975	0.499950	$2.5 \cdot 10^{-5}$	$2.5 \cdot 10^{-9}$

Más módszerrel (pl. KKT pont meghatározásán keresztül, a gyakorlás végett az olvasó oldja meg a feladatot így is) megoldva a feladatot, az optimális megoldás: $x_1 = 0.5$, $x_2 = 0.5$, $f_{\min} = 0.5$. A táblázatban jól látható, hogy a feladat megoldása hogyan konvergál az optimális megoldáshoz, továbbá az is látható, hogy a μ büntetőparaméter növekedésével a $\phi(\mathbf{x}_\mu)$, $f(\mathbf{x}_\mu)$ függvények növekszenek és az $f_{\min}$ értékhez tartanak, míg a $\mu B(\mathbf{x}_\mu)$, $B(\mathbf{x}_\mu)$ mennyiségek csökkennek és zérushoz tartanak.

4. Példa
Oldjuk meg az alábbi optimalizálási feladatot büntetőfüggvényes módszerrel:

$$\begin{aligned} f(x_1, x_2) &= (x_1 - 3)^2 + (x_2 - 2)^2 \to \min! \\ g_1(x_1, x_2) &= x_1^2 - x_2 - 3 \leqq 0 \\ g_2(x_1, x_2) &= x_1 + 2x_2 - 4 \leqq 0 \end{aligned}$$

Megoldás

A kiinduló vektor legyen $\mathbf{x} = (4,3)$, a kiinduló büntető paraméter pedig legyen $\mu = 1$. A feltétel nélküli segédfeladat:

$$\phi(\mathbf{x}) = (x_1-3)^2+(x_2-2)^2+\mu\cdot\left[\left(\max\left\{0, x_1^2 - x_2 - 3\right\}\right)^2 + (\max\{0, x_1 + 2x_2 - 4\})^2\right] \to \min!$$

A megoldás egyes lépéseit az alábbi táblázatban közöljük:

μ	$\mathbf{x}_\mu$	$\phi(\mathbf{x}_\mu)$	$f(\mathbf{x}_\mu)$	$\mu B(\mathbf{x}_\mu)$	$B(\mathbf{x}_\mu)$
1	$\begin{bmatrix} 2.074773 \\ 1.192522 \end{bmatrix}$	-11.267922	-11.491933	0.2240114	0.2240114
10	$\begin{bmatrix} 2.008504 \\ 1.022957 \end{bmatrix}$	-11.031470	-11.062322	0.0308523	0.0030852
100	$\begin{bmatrix} 2.0008625 \\ 1.0023407 \end{bmatrix}$	-11.003203	-11.006400	0.0031968	0.0000320
1000	$\begin{bmatrix} 2.0000864 \\ 1.0002343 \end{bmatrix}$	-11.000321	-11.000651	0.0003203	0.0000003

A fenti példát (a célfüggvény formájától eltekintve) egy másik tananyagban, ahol a KKT pont meghatározását tárgyaltuk, megoldottuk. Az optimális megoldás: $x_1 = 2$, $x_2 = 1$, $f_{\min} = -11$. A táblázatban jól látható a feladat megoldásának az optimális megoldáshoz való konvergálása, továbbá az is látható, hogy a μ büntetőparaméter növekedésével a $\phi(\mathbf{x}_\mu)$, $f(\mathbf{x}_\mu)$ függvények növekszenek és az $f_{\min}$ értékhez tartanak, míg a $\mu B(\mathbf{x}_\mu)$, $B(\mathbf{x}_\mu)$ mennyiségek csökkennek és zérushoz tartanak.

2.2. Korlátfüggvényes eljárás

A korlátfüggvényes eljárás gondolatmente hasonló a büntetőfüggvényes eljáráséhoz. Itt is beépítjük a feltételeket a célfüggvénybe és az így kapott feltétel nélküli optimalizálási feladatok sorozatával oldjuk meg az eredeti feltételes optimalizálási feladatot. A büntetőfüggvényes módszernél a megoldások fokozatosan elégítették ki a feltételeket, míg ennél az eljárásnál minden lépésben teljesednek a feltételek, azaz az eljárás során sohasem lépünk ki a megengedett (lehetséges) megoldások tartományából. Ezért ezt az eljárást szokás belsőpontos büntetőfüggvényes eljárásnak is nevezni. A belső pontban maradást úgy oldjuk meg, hogy ún. korlátfüggvény segítségével építjük be a feltételeket a célfüggvénybe. Ez a módszer megköveteli, hogy a feltételek között csak $g_i(\mathbf{x}) \leqq 0$ egyenlőtlenségi feltételek legyenek. A $K(\mathbf{x})$ korlátfüggvénynek általánosságban olyannak kell lenni, hogy a tartomány belsejében folytonos legyen és pozitív értéket adjon, míg a tartomány határához közeledve nagyon nagy legyen, tartson a $+\infty$-hez, képletben

$$K(y) = \begin{cases} \geqq 0, & \text{ha } y < 0 \\ \to +\infty, & \text{ha } y \to 0 \end{cases}$$

Összefoglalva tehát a következők mondhatók: Legyen a megoldandó feltételes optimalizálási feladat

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &\leqq 0 \quad i = 1, 2, \ldots, m \\ \mathbf{x} &\in X \end{aligned}$$

ehhez a leggyakrabban használatos $K(\mathbf{x})$ korlátfüggvények az alábbiak

$$K(\mathbf{x}) = \sum_{i=1}^{m} \frac{-1}{g_i(\mathbf{x})}, \quad \text{vagy} \quad K(\mathbf{x}) = -\sum_{i=1}^{m} \ln\left[\min\left\{1, -g_i(\mathbf{x})\right\}\right].$$

A második korlátfüggvényt ritkábban használják, mivel a $\min\{1, -g_i(\mathbf{x})\}$ függvény nem differenciálható. A gyakorlatban használatos az alábbi korlátfüggvény is, igaz, hogy a tulajdonságokat ez a függvény csak olyan $\mathbf{x}$ vektorok esetén teljesíti, amelyek a $g_i(\mathbf{x}) = 0$ kis környezetében vannak, így alkalmazása nem általános.

$$K(\mathbf{x}) = -\sum_{i=1}^{m} \ln\left[-g_i(\mathbf{x})\right].$$

A feladathoz tartozó segédfeladat

$$\begin{aligned} \phi(\mathbf{x}) = f(\mathbf{x}) + \mu K(\mathbf{x}) &\to \min! \\ g_i(\mathbf{x}) &< 0 \quad i = 1, 2, \ldots, m \\ \mathbf{x} &\in X \end{aligned}$$

A büntetőparaméter helyett itt korlátparamétert (μ) használunk, amely szintén pozitív, de az egyes segédfeladatokhoz tartozó értéke monoton csökkenően tart a zérushoz. A kiinduló vektornak ($\mathbf{x}_1$) belső pontnak ($\mathbf{x}_1 \in X,\ g_i(\mathbf{x}_1) \leqq 0, i = 1, 2, \ldots, m$) kell lennie, azaz első lépésként keresni kell egy lehetséges megoldást, ami sok esetben nem is olyan egyszerű feladat. Azon kívül, hogy az indulás sem olyan egyszerű ennél az eljárásnál, más nehézségek is akadnak. Először is az, hogy a segédfeladat látszólag nem feltétel nélküli. Ezen segíthetünk, mert a feltételeket elhagyhatjuk, ekkor viszont ellenőrizni kell, hogy a segédfeladat megoldása belső pont-e? Még ekkor sem egyszerű azonban a helyzet, mert a feltételek határának közelében nagy számok adódnak, amelyek számítástechnikai nehézségeket okozhatnak.

5. Példa

Az algoritmus illusztrálására oldjuk meg a 2. példában szereplő egyszerű optimalizálási feladatot korlátfüggvényes módszerrel

$$\begin{aligned} x &\to \min! \\ 1 - x &\leqq 0 \end{aligned}$$

Mint tudjuk, az optimális megoldás $\bar{x} = 1$, az optimális célfüggvényérték pedig $f(\bar{x}) = 1$. A feladathoz tartozó segédfeladat:

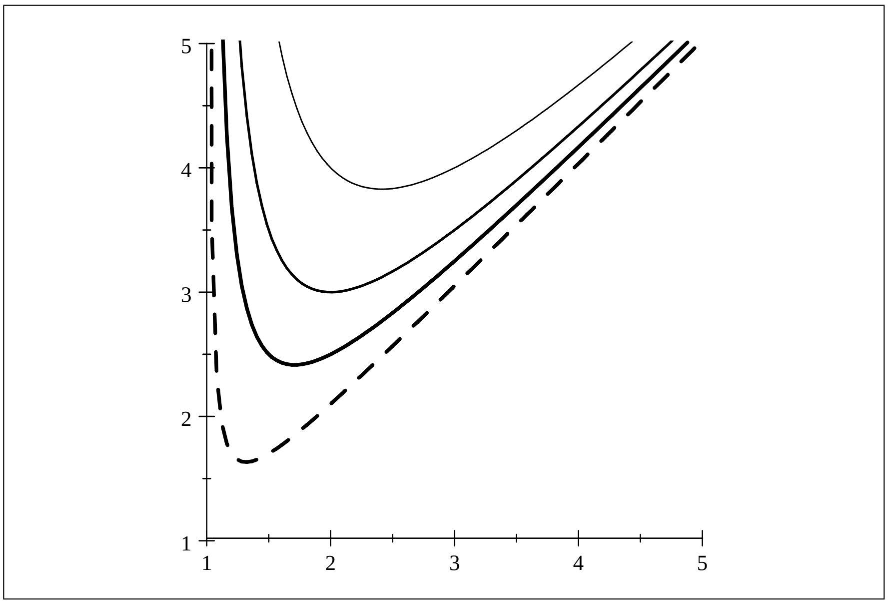
$$\phi(x) = f(x) + \mu \frac{-1}{1-x} \to \min!$$

amelynek optimális megoldása:

$$x_\mu = 1 + \sqrt{\mu}, \quad \mu > 0$$

Könnyen ellenőrizhető, hogy ha $\mu \to +0$, akkor az $x_\mu \to 1$, a feltételes optimalizálási feladat optimumához, $\bar{x}$-höz.

Az alábbi ábra segítségével szemléltetjük a korlátfüggvényes módszert. Az ábra a $\phi(x)$ segédfüggvényt ábrázolja $\mu = 2, \mu = 1, \mu = 0.5, \mu = 0.1$ korlátparaméterek esetén. A legvékonyabban rajzolt függvény a $\mu = 2$, a közepes vastagságú a $\mu = 1$, a vastag a $\mu = 0.5$, a szaggatottan rajzolt függvény pedig a $\mu = 0.1$ korlátparaméterhez tartozik.



Az ábrából nyomon követhetjük az alábbiakat: minél kisebb a μ korlátparaméter értéke, annál közelebb adódik a $\phi(x)$ segédfüggvény minimuma az eredeti feladat $f(x)$ célfüggvényének minimumához. Az ábra alapján az alábbi észrevételeket is tehetjük:

Ha $\mu \to 0$ akkor $x_\mu \to \bar{x} = 1$.
Ha $\mu \to 0$, akkor $f(x_\mu) \to f(\bar{x} = 1)$, μ csökkenésével monoton csökkenő módon.
Ha $\mu \to 0$, akkor $\phi(x_\mu) \to f(\bar{x} = 1)$, μ csökkenésével monoton csökkenő módon.
Ha $\mu \to 0$, akkor $\mu B(x_\mu) \to 0$, μ csökkenésével monoton csökkenő módon.
Ha $\mu \to 0$, akkor $B(x_\mu) \to 0$, μ csökkenésével monoton csökkenő módon.

A következő tétel a fenti észrevételek pontosítását adja:

TÉTEL
Tekintsük az alábbi optimalizálási feladat

$$\begin{array}{ll} f(\mathbf{x}) \to \min! & \\ g_i(\mathbf{x}) \leqq 0 & i = 1, 2, \ldots, m \\ \mathbf{x} \in X & \end{array}$$

ahol f, g_i $i = 1, 2, \ldots, m$ folytonos függvények $\mathbb{R}^n$-ben, $X \subseteq \mathbb{R}^n$ nemüres zárt halmaz. Tegyük fel, hogy az $\{\mathbf{x} \in X : g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\}$ halmaz nem üres és a $K(\mathbf{x})$ korlátfüggvény folytonos az $\{\mathbf{x} \in X : g_i(\mathbf{x}) \leqq 0, i = 1, 2, \ldots, m\}$ halmazon. Továbbá tegyük fel, hogy minden nemnegatív μ skalárra létezik $\mathbf{x}_\mu$ optimális megoldása a

$$\min \{\phi(\mathbf{x}) = f(\mathbf{x}) + \mu K(\mathbf{x}) : \mathbf{x} \in X, g_i(\mathbf{x}) \leq 0\}$$

optimalizálási feladatnak azzal a tulajdonsággal, hogy az $\mathbf{x}_\mu$ minden $N(\mathbf{x}_\mu)$ környezetében van olyan $\mathbf{x} \in X \cap N(\mathbf{x}_\mu)$, hogy $g_i(\mathbf{x}) < 0, i = 1, 2, \ldots, m$. Ekkor

$$\begin{aligned} & \min \{f(\mathbf{x}) : g_i(\mathbf{x}) \leqq 0, (i = 1, 2, \ldots, m), \mathbf{x} \in X\} = \\ = \ & \lim_{\mu \to 0+} \Theta(\mu) = \inf \{\Theta(\mu) : \mu > 0\}, \end{aligned}$$

ahol $\Theta(\mu) = f(\mathbf{x}_\mu) + \mu K(\mathbf{x}_\mu)$. Továbbá az $\{\mathbf{x}_\mu\}$ sorozat bármely konvergens részsorozata tart az $\bar{\mathbf{x}}$ ponthoz, amely az eredeti optimalizálási feladat optimális megoldása. Ezenfelül, ha $\mu \to 0+$, akkor igazak az alábbiak:

a) $f(\mathbf{x}_\mu) \to f(\bar{\mathbf{x}})$ monoton csökkenően,
b) $\phi(\mathbf{x}_\mu) \to f(\bar{\mathbf{x}})$ monoton csökkenően,
c) $\Theta(\mu) \to f(\bar{\mathbf{x}})$ monoton csökkenően,
d) $\mu K(\mathbf{x}_\mu) \to 0$ monoton csökkenően,
e) $K(\mathbf{x}_\mu) \to 0$ monoton csökkenően.

Az alábbiakban a korlátfüggvényes módszer megoldási módját írjuk le.

Korlátfüggvényes eljárás algoritmusa:
Induló lépés ($k = 1$): Kiindulunk egy olyan $\mathbf{x}_k \in \mathbb{R}^n$ vektorból, amelyre $g_i(\mathbf{x}_k) < 0$ minden i indexre, azaz egy belső pontból, választunk egy $\mu_k > 0$ korlátparamétert és egy $0 < \beta < 1$ skalárt.
Közbülső lépés:
1. Megoldjuk a $\phi(\mathbf{x}) = \min\{f(\mathbf{x}) + \mu_k K(\mathbf{x}) : \mathbf{x} \in X, g_i(\mathbf{x}) < 0\}$ minimalizálási feladatot, legyen az optimális megoldás $\mathbf{x}_{k+1}$.
2. Megállunk, ha $\mu_k K(\mathbf{x}_{k+1}) < \varepsilon$, egyébként folytatjuk az eljárást következő sorral.
3. Meghatározzuk a korlátparaméter következő értékét, legyen $\mu_{k+1} = \beta\mu_k$.
4. Folytatjuk az eljárást, $k := k + 1$.

3. Lehetséges irányok módszere

A feltétel nélküli optimalizálási feladat megoldásában megismert vonalmenti optimalizálási módszerek a következő elven működtek: kiindultunk **egy tetszőleges pontból**, majd valamilyen módszerrel meghatároztunk **egy irányt** és abban az irányban megkerestük azt a pontot, ahol a célfüggvény értéke minimális. A feltételes optimalizálási feladat megoldásában a most ismertetésre kerülő lehetséges irányok módszere hasonló elven működik: kiindulunk **egy lehetséges (megengedett) megoldásból**, valamilyen módszerrel meghatározunk egy olyan irányt, amely egyúttal **javító (csökkenő) és lehetséges (megengedett) irány** és ebben az irányban megkeressük a célfüggvény minimumát ügyelve arra, hogy a lehetséges megoldások tartományában maradjunk.

3.1. Javító lehetséges irány

Ismétlésképpen újra megfogalmazzuk a lehetséges irány és a javító irány fogalmakat.
Tekintsük a megszokott $\min\{f(\mathbf{x}) : \mathbf{x} \in S \subset \mathbb{R}^n\}$ feltételes optimalizálási feladatot.

Lehetséges irány definíciója
Egy nemzérus $\mathbf{d}$ vektort az $\mathbf{x} \in S$ pontban lehetséges iránynak nevezünk, ha létezik olyan $\delta > 0$ szám, hogy
$$\mathbf{x} + \lambda\mathbf{d} \in S \quad \text{minden } \lambda \in (0, \delta)\text{-ra.}$$

Javító irány definíciója
Egy nemzérus $\mathbf{d}$ vektort az $\mathbf{x} \in S$ pontban javító iránynak nevezünk, ha létezik olyan $\delta > 0$ szám, hogy
$$f(\mathbf{x} + \lambda\mathbf{d}) < f(\mathbf{x}) \quad \text{minden } \lambda \in (0, \delta)\text{-ra.}$$

A két fogalom egyesítéséből a javító lehetséges irányt az alábbiakban fogalmazhatjuk meg.

Javító lehetséges irány definíciója
Egy nemzérus $\mathbf{d}$ vektort az $\mathbf{x} \in S$ pontban javító lehetséges iránynak nevezünk, ha létezik olyan $\delta > 0$ szám, hogy

$$f(\mathbf{x} + \lambda\mathbf{d}) < f(\mathbf{x}) \text{ és } \mathbf{x} + \lambda\mathbf{d} \in S \quad \text{minden } \lambda \in (0, \delta)\text{-ra.}$$

Javító lehetséges irány karakterizációja
Amennyiben a célfüggvény és a feltételi halmazt leíró függvények differenciálhatók, akkor a javító lehetséges irányt a gradiens segítségével is meghatározhatjuk, ezt fejezi ki az alábbi karakterizáció:

Legyen $S = \{\mathbf{x} : \mathbf{x} \in X, \quad g_i(\mathbf{x}) \leqq 0, \ i = 1, 2, \ldots, m\}$, ahol $g_1, g_2, \ldots, g_m : R^n \to R$ függvények és $X \subseteq R^n$ nemüres nyílt halmaz. Legyen adott egy $\mathbf{x} \in S$ vektor és legyen $I = \{i : g_i(\mathbf{x}) = 0\}$ az aktív feltételi függvények indexhalmaza. Tegyük fel, hogy a g_i $i \in I$ feltételi függvények, az f célfüggvény az $\mathbf{x}$ pontban differenciálhatók, a $g_i, i \notin I$ függvények pedig az $\mathbf{x}$ pontban folytonosak. Ekkor a

$$\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} < 0, \ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d} < 0, \ i \in I$$

feltételeket kielégítő $\mathbf{d}$ vektorok javító lehetséges irányvektorok. Ennek a karakterizációnak a bizonyítása egy másik

A módszerek tárgyalásánál látni fogjuk, hogy a javító lehetséges irányt más-más elven fogjuk meghatározni. A javító lehetséges irányt az első módszernél egy lineáris programozási feladat megoldása szolgáltatja, a második módszerben pedig a legjobb javító irány vetítéseként kapjuk.

3.2. Zoutendijk módszere

Először olyan feltételes optimalizálási feladatokat vizsgálunk, amelyekben a feltételi függvények lineárisak.

3.2.1. Lineáris feltételek esete

A feltételes optimalizálási feladat általános alakja lineáris feltételek esetén az alábbi:

$$\begin{aligned} f(\mathbf{x}) &\to \min! \\ \mathbf{A}\mathbf{x} &\leqq \mathbf{b} \\ \mathbf{B}\mathbf{x} &= \mathbf{c} \end{aligned}$$

ahol $f : \mathbb{R}^n \to \mathbb{R}, \mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{k\times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^k$.

TÉTEL
Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennáll, hogy $\mathbf{A}\mathbf{x} \leqq \mathbf{b}, \ \mathbf{B}\mathbf{x} = \mathbf{c}$. Az $\mathbf{x}$ megoldásnál az egyenlőtlenséges feltételek egyik része egyenlőség, míg a másik része szigorú egyenlőtlenség

formájában teljesedik. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x} < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói.

a) Egy nemzérus $\mathbf{d}$ vektor akkor és csak akkor lehetséges irány, ha

$$\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}, \quad \mathbf{Bd} = \mathbf{0}.$$

b) Ha $\nabla f(\mathbf{x})\mathbf{d} < 0$, akkor $\mathbf{d}$ egy javító irány.

Bizonyítás

Ha $\mathbf{x}$ egy lehetséges megoldás, ekkor az $\mathbf{x} + \lambda\mathbf{d}$ vektor akkor lesz valamilyen pozitív λ-ra lehetséges megoldás, ha

$$\begin{aligned} \mathbf{A}(\mathbf{x} + \lambda\mathbf{d}) &\leqq \mathbf{b}, \\ \mathbf{A}(\mathbf{x} + \lambda\mathbf{d}) &= \mathbf{c}. \end{aligned}$$

Részletezve kapjuk, hogy

$\mathbf{A}_1(\mathbf{x} + \lambda\mathbf{d}) = \mathbf{A}_1\mathbf{x} + \lambda\mathbf{A}_1\mathbf{d} \leqq \mathbf{b}_1$, mivel $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$, így az egyenlőtlenség akkor és csak akkor teljesedhet valamilyen pozitív λ-ra, ha $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$,

$\mathbf{A}_2(\mathbf{x}+\lambda\mathbf{d}) = \mathbf{A}_2\mathbf{x}+\lambda\mathbf{A}_2\mathbf{d} \leqq \mathbf{b}_2$, mivel $\mathbf{A}_2\mathbf{x} < \mathbf{b}_2$, így az egyenlőtlenség mindig teljesedik valamilyen pozitív λ-ra,

$\mathbf{B}(\mathbf{x} + \lambda\mathbf{d}) = \mathbf{Bx} + \lambda\mathbf{Bd} = \mathbf{c}$, mivel $\mathbf{Bx} = \mathbf{c}$, így az egyenlőség akkor és csak akkor teljesedhet valamilyen pozitív λ-ra, ha $\mathbf{Bd} = \mathbf{0}$.

A $\nabla f(\mathbf{x})\mathbf{d}$ mennyiség a $\mathbf{d}$ irányban az iránymenti derivált, ha ez negatív, akkor ebben az irányban a célfüggvény csökken, tehát $\mathbf{d}$ javító irány. **Q.e.d.**

A tétel szerint tehát egy javító lehetséges irányra fenn kell állnia, hogy

$$\nabla f(\mathbf{x})\mathbf{d} < 0, \quad \mathbf{A}_1\mathbf{d} \leqq \mathbf{0}, \quad \mathbf{Bd} = \mathbf{0}.$$

Természetes módon adódik, hogy azt a $\mathbf{d}$ irányt keresssük, amely a legjobb javító irány a lehetséges irányok között. Ilyen irányt az alábbi lineáris programozási feladat (LP) megoldásaként nyerhetünk:

$$\begin{aligned} \nabla f(\mathbf{x})\mathbf{d} &\rightarrow \min! \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{Bd} &= \mathbf{0}. \end{aligned}$$

Könnyen látható azonban, hogy ezen LP feladat feltételi halmaza nem korlátos, ha ugyanis egy $\mathbf{d}$ vektor megoldás, akkor $\lambda\mathbf{d}$ is megoldás minden $\lambda > 0$ esetén, így a célfüggvény nem véges, azaz $\nabla f(\mathbf{x})\mathbf{d} \rightarrow -\infty$. Ezen többféleképpen is segíthetünk ún. normalizáló feltétel használatával. Nevezetesen vagy a $\mathbf{d}$ vektor nagyságára írunk elő egy felső korlátot vagy a célfüggvényre írunk elő egy alsó korlátot. Az alábbi három feladatot szoktuk alkalmazni, amelyek a következők.

I. feladat

$$\begin{aligned} \nabla f(\mathbf{x})\mathbf{d} &\rightarrow \min! \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{Bd} &= \mathbf{0}, \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \ldots, n \end{aligned}$$

II. feladat

$$\begin{aligned} \nabla f(\mathbf{x})\mathbf{d} &\rightarrow \text{min!} \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{B}\mathbf{d} &= \mathbf{0}, \\ \mathbf{d}^2 &\leqq 1. \end{aligned}$$

III. feladat

$$\begin{aligned} \nabla f(\mathbf{x})\mathbf{d} &\rightarrow \text{min!} \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{B}\mathbf{d} &= \mathbf{0}, \\ \nabla f(\mathbf{x})\mathbf{d} &\geqq -1. \end{aligned}$$

Az első és a harmadik feladat lineáris programozási (LP) feladat, míg a második egy kvadratikus feltétellel rendelkezik. Mivel mindegyik feladatnak a $\mathbf{d} = \mathbf{0}$ lehetséges megoldása és ekkor a célfüggvény értéke zérus, ebből következik, hogy a célfüggvény optimális értéke nem lehet pozitív.

TÉTEL

Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \text{min!} \\ \mathbf{A}\mathbf{x} &\leqq \mathbf{b} \\ \mathbf{B}\mathbf{x} &= \mathbf{c} \end{aligned}$$

Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennáll, hogy $\mathbf{A}\mathbf{x} \leqq \mathbf{b}, \mathbf{B}\mathbf{x} = \mathbf{c}$. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x} < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói.

Ekkor az $\mathbf{x}$ lehetséges megoldás akkor és csak akkor Karush-Kuhn-Tucker (KKT) pont, ha a I., II., III. feladatok célfüggvényének optimális (minimális) értéke zérus.

Bizonyítás

A Karush-Kuhn-Tucker feltételek vizsgálatánál láttuk, hogy egy $\mathbf{x}$ vektor akkor és csak akkor KKT pont, ha létezik olyan $\mathbf{u} \geqq \mathbf{0}$ és $\mathbf{v}$ vektor, hogy

$$\nabla f(\mathbf{x}) + \mathbf{A}_1^T\mathbf{u} + \mathbf{B}^T\mathbf{v} = \mathbf{0}.$$

Az alapfeladatban a $g_i(\mathbf{x}) \leqq 0$ és a $h_i(\mathbf{x}) = 0$ feltételek vektoriálisan a $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ és a $\mathbf{h}(\mathbf{x}) = \mathbf{B}\mathbf{x}$ alakban írhatók. Ezen függvények gradiensei pedig a következők: $\nabla\mathbf{g} = \nabla(\mathbf{A}_1\mathbf{x} - \mathbf{b}) = \mathbf{A}_1^T$, $\nabla\mathbf{h} = \nabla(\mathbf{B}\mathbf{x} - \mathbf{c}) = \mathbf{B}^T$. A KKT feltételben pedig ezen gradiens vektorok lineáris kombinációja szerepel. Most tekintsük a Farkas tételnek idevágó változatát (egy másik tananyagban ismertettük), amely az alábbi két rendszer megoldhatóságára vonatkozik:

$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} = \mathbf{b}$ $\mathbf{x} \geqq \mathbf{0}$	$\mathbf{y}\mathbf{A} \leqq \mathbf{0}$ $\mathbf{y}\mathbf{B} = \mathbf{0}$ $\mathbf{y}\mathbf{b} > 0$

A mi esetünkben a baloldali rendszer olyan, hogy $\mathbf{A} = \mathbf{A}_1^T$, $\mathbf{B} = \mathbf{B}^T$, $\mathbf{x} = \mathbf{u}$, $\mathbf{z} = \mathbf{v}$, $\mathbf{b} = -\boldsymbol{\nabla} f(\mathbf{x})$, amelynek sémája az alábbi:

	$\mathbf{u}$	$\mathbf{v}$	
$\mathbf{d}$	$\mathbf{A}^T$	$\mathbf{B}^T$	$-\boldsymbol{\nabla} f(\mathbf{x})$

A Farkas tétel szerint a fenti két rendszer közül egyik és csak egyik oldható meg. Eszerint akkor és csak akkor léteznek $\mathbf{u}$ és $\mathbf{v}$ vektorok, ha a

$$\begin{aligned} \boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} &< 0, \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{B}\mathbf{d} &= \mathbf{0}, \end{aligned}$$

rendszer nem oldható meg. Az $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$, $\mathbf{B}\mathbf{d} = \mathbf{0}$ rendszer megoldható, hisz a $\mathbf{d} = \mathbf{0}$ megoldás. A teljes rendszer akkor és csak akkor nem oldható meg, ha $\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} \geqq 0$. Korábbról láttuk, hogy a $\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d}$ célfüggvény optimális értéke $\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} > 0$ nem lehet egyik feladatnál sem, így az optimális értékre $\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} = 0$ adódik. **Q.e.d.**

Iránykeresés

Mivel az alapfeladatunk egy feltételes optimalizálási feladat, így a javító lehetséges irányban egy pont meghatározása szintén feltételes optimalizálási feladat lesz, igaz, hogy egyváltozós optimalizálási feladatot kell megoldanunk. Legyen adott egy $\mathbf{x}_k$ lehetséges megoldás és a hozzátartozó $\mathbf{d}_k$ javító lehetséges irány. A λ meghatározására az alábbi egyváltozós optimalizálási feladat szolgál:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda\mathbf{d}_k) \to \min! \\ \mathbf{A}(\mathbf{x}_k + \lambda\mathbf{d}_k) &\leqq \mathbf{b} \\ \mathbf{B}(\mathbf{x}_k + \lambda\mathbf{d}_k) &= \mathbf{c} \\ \lambda &\geqq 0 \end{aligned}$$

Tegyük fel, hogy $\mathbf{A}_1\mathbf{x}_k = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x}_k < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói. Az aktív egyenlőtlenségekből és az egyenletekből nem kapunk semmiféle megszorítást λ-ra, hiszen

$\mathbf{A}_1(\mathbf{x}_k + \lambda\mathbf{d}_k) = \mathbf{A}_1\mathbf{x}_k + \lambda\mathbf{A}_1\mathbf{d}_k \leqq \mathbf{b}_1$, mivel $\mathbf{A}_1\mathbf{x}_k = \mathbf{b}_1$ és $\mathbf{A}_1\mathbf{d}_k \leqq \mathbf{0}$, így λ bármilyen pozítív szám lehet,

$\mathbf{B}(\mathbf{x}_k + \lambda\mathbf{d}_k) = \mathbf{B}\mathbf{x}_k + \lambda\mathbf{B}\mathbf{d}_k = \mathbf{c}$, mivel $\mathbf{B}\mathbf{x}_k = \mathbf{c}$ és $\mathbf{B}\mathbf{d}_k = \mathbf{0}$, így λ bármilyen pozítív szám lehet.

Az inaktív feltételekből viszont már λ-ra kapunk előírásokat, amelyek a következők szerint alakulnak:

$$\mathbf{A}_2(\mathbf{x}_k + \lambda\mathbf{d}_k) = \mathbf{A}_2\mathbf{x}_k + \lambda\mathbf{A}_2\mathbf{d}_k \leqq \mathbf{b}_2, \text{ amelyből } \lambda\mathbf{A}_2\mathbf{d}_k \leqq \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k.$$

Ha az $\mathbf{A}_2\mathbf{d}_k$ vektor valamely eleme nempozitív, akkor λ bármilyen pozítív szám lehet. Jelöljük egy vektor i-edik elemét a $(\ldots)_i$ szimbólummal. Ha $(\mathbf{A}_2\mathbf{d}_k)_i > 0$, akkor

$$\lambda \leqq \frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i},$$

így λ maximális értéke

$$\lambda_{\max} = \min\left\{\frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i} : (\mathbf{A}_2\mathbf{d}_k)_i > 0\right\}$$

Mivel $\mathbf{A}_2\mathbf{x}_k < \mathbf{b}_2$, így ha van legalább egy $(\mathbf{A}_2\mathbf{d}_k)_i > 0$ elem, akkor $\lambda_{\max} > 0$, amennyiben $\mathbf{A}_2\mathbf{d}_k \leqq \mathbf{0}$, úgy λ-ra az inaktív feltételek sem adnak korlátozást, ekkor $\lambda_{\max} = \infty$.

Összefoglalva az iránymenti optimalizálási feladat a következő:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda\mathbf{d}_k) \to \min! \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

ahol

$$\lambda_{\max} = \begin{cases} \min\left\{\frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i} : (\mathbf{A}_2\mathbf{d}_k)_i > 0\right\}, & \text{ha } \mathbf{A}_2\mathbf{d}_k \nleqq \mathbf{0} \\ \infty, & \text{ha } \mathbf{A}_2\mathbf{d}_k \leqq \mathbf{0} \end{cases}$$

Zoutendijk módszer algoritmusa lineáris feltételek esetén:

Induló lépés ($k = 1$): Kiindulunk egy $\mathbf{x}_k$ lehetséges megoldásból, amelyre $\mathbf{A}\mathbf{x}_k \leqq \mathbf{b}, \mathbf{B}\mathbf{x}_k = \mathbf{c}$.

Közbülső lépés:

1. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x}_k = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x}_k < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói.

2. Meghatározzuk az $\mathbf{x}_k$ lehetséges megoldáshoz a javító lehetséges irányt, azaz megoldjuk az alábbi lineáris programozási feladatot:

$$\begin{aligned} \boldsymbol{\nabla} f(\mathbf{x}_k)\mathbf{d} &\to \min! \\ \mathbf{A}_1\mathbf{d} &\leqq \mathbf{0}, \\ \mathbf{B}\mathbf{d} &= \mathbf{0}, \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \ldots, n \end{aligned}$$

   Legyen az optimális megoldás $\mathbf{d}_k$. (Megjegyezzük, hogy a II. és a III. feladatban használt normalizáló feltételt is használhattuk volna.)

3. Megállunk, ha az optimális célfüggvény érték $\min(\boldsymbol{\nabla} f(\mathbf{x}_k)\mathbf{d}_k) = 0$. Ekkor az $\mathbf{x}_k$ lehetséges megoldás KKT pont. Egyébként folytatjuk az eljárást a következő sorral.

4. Megoldjuk a $\mathbf{d}_k$ javító lehetséges irányban az iránymenti optimalizálási feladatot, amely az alábbi:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda\mathbf{d}_k) \to \min! \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

   ahol

$$\lambda_{\max} = \begin{cases} \min\left\{\frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i} : (\mathbf{A}_2\mathbf{d}_k)_i > 0\right\}, & \text{ha } \mathbf{A}_2\mathbf{d}_k \nleqq \mathbf{0} \\ \infty, & \text{ha } \mathbf{A}_2\mathbf{d}_k \leqq \mathbf{0} \end{cases}$$

   Legyen az optimális megoldás λ_k.

5. Meghatározzuk a következő közelítést: $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k$.

6. Folytatjuk az eljárást, $k := k + 1$.

6. Példa

Legyenek adottak az $x_1 - x_2 = 2$ és az $x_1 + x_2 = 4$ egyenesek. A két egyenes az (x_1, x_2) síkot négy tartományra osztja. Tekintsük azt a tartományt, amely a $(7, 7)$ pontot tartalmazza. Határozzuk meg a tartomány azon pontját, amely az origóhoz a legközelebb van! Oldjuk meg az optimalizálási feladatot Zoutendijk módszerrel!

Megoldás:

Először felírjuk a feladatot matematikai formában. Az egyenesek egyenletébe behelyettesítve a pontot, megkapjuk a feltételeket. Az első feltétel $x_1 - x_2 \leqq 2$, a második pedig $x_1 + x_2 \geqq 4$, amelynek szükséges alakja $-x_1 - x_2 \leqq -4$. A távolság helyett annak négyzetét használhatjuk, így a célfüggvény $x_1^2 + x_2^2$. A megoldandó optimalizálási feladat tehát a következő:

$$
\begin{aligned}
x_1^2 + x_2^2 &\rightarrow \text{min!} \\
x_1 - x_2 &\leqq 2 \\
-x_1 - x_2 &\leqq -4
\end{aligned}
$$

A célfüggvény gradiense: $\boldsymbol{\nabla} f(\mathbf{x}) = (2x_1, 2x_2)$.

$k = 1$

Legyen az $\mathbf{x}_1 = (5, 3)$ egy lehetséges megoldás. Az első feltétel aktív, a második feltétel inaktív. A célfüggvény gradiense az $\mathbf{x}_1 = (5, 3)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_1) = (10, 6)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő (célfüggvény értékére tettünk normáló feltételt):

$$
\begin{aligned}
10d_1 + 6d_2 &\rightarrow \text{min!} \\
d_1 - d_2 &\leqq 0 \\
10d_1 + 6d_2 &\geqq -1
\end{aligned}
$$

A fenti lineáris programozási feladat optimális megoldása: $d_1 = -1/10, d_2 = 0$, azaz az $\mathbf{x}_1 = (5, 3)$ ponthoz tartozó $\mathbf{d}_1$ irányvektor vektorosan: $\mathbf{d}_1 = (-1/10, 0)$. A célfüggvény minimális értéke: -1, mivel nem zérus, így az $\mathbf{x}_1 = (5, 3)$ megoldás nem KKT pont.

A fenti lineáris programozási feladat kétváltozós, így grafikusan is megoldható. Nagyobb méreteknél azonban szimplex módszerrel oldjuk meg az LP feladatot, de ekkor az LP feladatban szereplő változóknak nemnegatiíaknak kell lenni. A d_1, d_2 változók helyett (mivel ezek lehetnek negatívok is), új nemnegativ változókat vezetünk be az alábbiak szerint. Legyen $d_1 = d_1^+ - d_1^-$, ill. $d_2 = d_2^+ - d_2^-$, ahol $d_1^+, d_1^-, d_2^+, d_2^- \geqq 0$. Így a szimplex módszerrel már megoldható lineáris programozási feladat az alábbi

$$
\begin{aligned}
10d_1^+ - 10d_1^- + 6d_2^+ - 6d_2^- &\rightarrow \text{min!} \\
d_1^+ - d_1^- - d_2^+ + d_2^- &\leqq 0 \\
-10d_1^+ + 10d_1^- - 6d_2^+ + 6d_2^- &\leqq 1 \\
d_1^+, d_1^-, d_2^+, d_2^- &\geqq 0
\end{aligned}
$$

Javasoljuk, hogy a könnyebb számolás kedvéért ne a törtértékű $\mathbf{d}_1 = (-1/10, 0)$ irányvektorral dolgozzunk, hanem a $\mathbf{d}_1 = (-1, 0)$ irányvektorral, mivel csak az irány az érdekes az iránymenti optimalizálásban. Az új $\mathbf{x}$ vektor a $\mathbf{d}_1$ irányban $\mathbf{x} = \mathbf{x}_1 + \lambda\mathbf{d}_1 = (5,3) + \lambda(-1,0) = (5-\lambda, 3)$. A $\lambda_{\max}$ értékének meghatározása

$$\lambda_{\max} = \min\left\{\frac{-4-(-5-3)}{1}\right\} = 4.$$

Természetesen képlet nélkül is meghatározhatjuk $\lambda_{\max}$ értékét, nem kell mást tennünk, mint az inaktív feltételekbe behelyettesíteni az $\mathbf{x} = (5-\lambda, 3)$ vektort és meghatározni azt a λ-t, amelynél még érvényesek a feltételek, jelen példában

$$-(5-\lambda) - 3 \leqq -4,$$

amelyből $\lambda \leqq 4$, így $\lambda_{\max} = 4$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (5-\lambda)^2 + 3^2 \to \min! \\ 0 &\leqq \lambda \leqq 4 \end{aligned}$$

Mivel $\varphi(\lambda)$ kvadratikus függvény, így nem kell különösebb eszköz az optimalizáláshoz, az optimális megoldás $\lambda_1 = 4$, így a következő pont $\mathbf{x}_2 = \mathbf{x}_1 + \lambda_1\mathbf{d}_1 = (5,3) + 4(-1,0) = (1,3)$.

$k = 2$

Az $\mathbf{x}_2 = (1,3)$ lehetséges megoldásnál az első feltétel inaktív, a második feltétel aktív. A célfüggvény gradiense az $\mathbf{x}_2 = (1,3)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_2) = (2,6)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő (célfüggvény értékére tettünk normáló feltételt):

$$\begin{aligned} 2d_1 + 6d_2 &\to \min! \\ -d_1 - d_2 &\leqq 0 \\ 2d_1 + 6d_2 &\geqq -1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása, azaz a következő irányvektor vektorosan: $\mathbf{d}_2 = (1,-1)$. A célfüggvény minimális értéke: -2, mivel nem zérus így az $\mathbf{x}_1 = (1,3)$ megoldás nem KKT pont.

A $\mathbf{d}_2 = (1,-1)$ irányban az új $\mathbf{x}$ vektor $\mathbf{x} = \mathbf{x}_2 + \lambda\mathbf{d}_2 = (1,3) + \lambda(1,-1) = (1+\lambda, 3-\lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\lambda_{\max} = \min\left\{\frac{2-(1-3)}{2}\right\} = 2.$$

Vagy képlet nélkül az inaktív feltételekbe behelyettesítjük az $\mathbf{x} = (1+\lambda, 3-\lambda)$ vektort , azaz

$$(1+\lambda) - (3-\lambda) \leqq 2,$$

amelyből $\lambda \leqq 2$, így $\lambda_{\max} = 2$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (1+\lambda)^2 + (3-\lambda)^2 \to \min! \\ 0 &\leqq \lambda \leqq 2 \end{aligned}$$

Ennek optimális megoldása $\lambda_2 = 1$, így a következő pont $\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2\mathbf{d}_2 = (2,2)$.

$k = 3$

Az $\mathbf{x}_3 = (2,2)$ lehetséges megoldásnál az első feltétel inaktív, a második feltétel aktív. A célfüggvény gradiense az $\mathbf{x}_3 = (2,2)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_3) = (4,4)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő (célfüggvény értékére tettünk normáló feltételt):

$$\begin{aligned} 4d_1 + 4d_2 &\rightarrow \text{min!} \\ -d_1 - d_2 &\leqq 0 \\ 4d_1 + 4d_2 &\geqq -1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $\mathbf{d}_3 = (0,0)$ és a minimális célfüggvényérték $\min(\boldsymbol{\nabla} f(\mathbf{x}_3)\mathbf{d}_3) = 0$. Megállunk, mivel az optimális célfüggvényérték zérus, így a tétel értelmében az $\mathbf{x}_3 = (2,2)$ lehetséges megoldás KKT pont.

7. Példa

Tekintsük az előző példabeli feladatot, de most az iránykeresésnél az irányvektor koordinátáira tegyünk normáló feltételeket.

$k = 1$

Legyen itt is az $\mathbf{x}_1 = (5,3)$ egy lehetséges megoldás. Az első feltétel aktív, a második feltétel inaktív. A célfüggvény gradiense az $\mathbf{x}_1 = (5,3)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_1) = (10,6)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő:

$$\begin{aligned} 10d_1 + 6d_2 &\rightarrow \text{min!} \\ d_1 - d_2 &\leqq 0 \\ -1 &\leqq d_1 \leqq 1 \\ -1 &\leqq d_2 \leqq 1 \end{aligned}$$

A fenti lineáris programozási feladat szimplex módszerrel való megoldásához az új nemnegatív változók bevezetése ($d_1 = d_1^+ - d_1^-$, $d_2 = d_2^+ - d_2^-$,) után az alábbi lineáris programozási feladatot kell megoldani:

$$\begin{aligned} 10d_1^+ - 10d_1^- + 6d_2^+ - 6d_2^- &\rightarrow \text{min!} \\ d_1^+ - d_1^- - d_2^+ + d_2^- &\leqq 0 \\ d_1^+ - d_1^- &\leqq 1 \\ d_1^- - d_1^+ &\leqq 1 \\ d_2^+ - d_2^- &\leqq 1 \\ d_2^- - d_2^- &\leqq 1 \\ d_1^+, d_1^-, d_2^+, d_2^- &\geqq 0 \end{aligned}$$

Megjegyzés:

Láthatjuk, hogy a változók száma megduplázódott és a feltételek száma is megnőtt, van olyan lehetőség is a nemnegativitás biztosítására, hogy a változók száma nem változzon és a feltételek száma is csak mérsékeltebben növekedjen.

Ha a $-1 \leqq d_i \leqq 1$ $(i = 1,2)$ előírásokat a $0 \leqq d_i + 1 \leqq 2$ $(i = 1,2)$ alakban írjuk fel, akkor kézenfekvő az olyan új változók $(\bar{d}_1, \bar{d}_2)$ bevezetése, amelyek az alábbiak: $\bar{d}_1 = d_1 + 1$, $\bar{d}_2 = d_2 + 1$. Ekkor a $0 \leqq \bar{d}_i \leqq 2$ $(i = 1,2)$ előírások adódnak, amelyből a $\bar{d}_1, \bar{d}_2 \geqq 0$ nemnegativitási feltétel nem ad új feltételt, csak a $\bar{d}_1, \bar{d}_2 \leqq 2$ új feltételeket kell felvenni. Ezzel a módszerrel az irányvektorok meghatározására az alábbi lineáris programozási feladatot kell megoldani:

$$\begin{aligned} 10(\bar{d}_1 - 1) + 6(\bar{d}_2 - 1) &\rightarrow \text{min!} \\ (\bar{d}_1 - 1) - (\bar{d}_2 - 1) &\leqq 0 \\ \bar{d}_1 &\leqq 2 \\ \bar{d}_2 &\leqq 2 \\ \bar{d}_1, \bar{d}_2 &\geqq 0 \end{aligned}$$

A feladatot szimplex módszerrel oldjuk meg. Az induló szimplex tábla az alábbi:

	$\bar{d}_1$	$\bar{d}_2$	
u_1	1	–1	0
u_2	1	0	2
u_2	0	1	2
	-10	-6	-16

Az induló szimplex tábla optimális, az optimális megoldás: $\bar{d}_1 = 0$, $\bar{d}_2 = 0$. A következő irányvektor koordinátái: $d_1 = -1, d_2 = -1$, vektorosan: $\mathbf{d}_1 = (-1, -1)$. A célfüggvény minimális értéke: -16, mivel nem zérus, így az $\mathbf{x}_1 = (5, 3)$ megoldás nem KKT pont.

A $\mathbf{d}_1 = (-1, -1)$ irányban az új $\mathbf{x}$ vektor $\mathbf{x} = \mathbf{x}_2 + \lambda d_2 = (5 - \lambda, 3 - \lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\lambda_{\max} = \min \left\{ \frac{-4 - (-5 - 3)}{2} \right\} = 2.$$

Vagy képlet nélkül, az inaktív feltételekbe behelyettesítve az $\mathbf{x} = (5 - \lambda, 3 - \lambda)$ vektor, kapjuk, hogy és meghatározni azt a λ-t, amelynél még érvényesek a feltételek, jelen példában

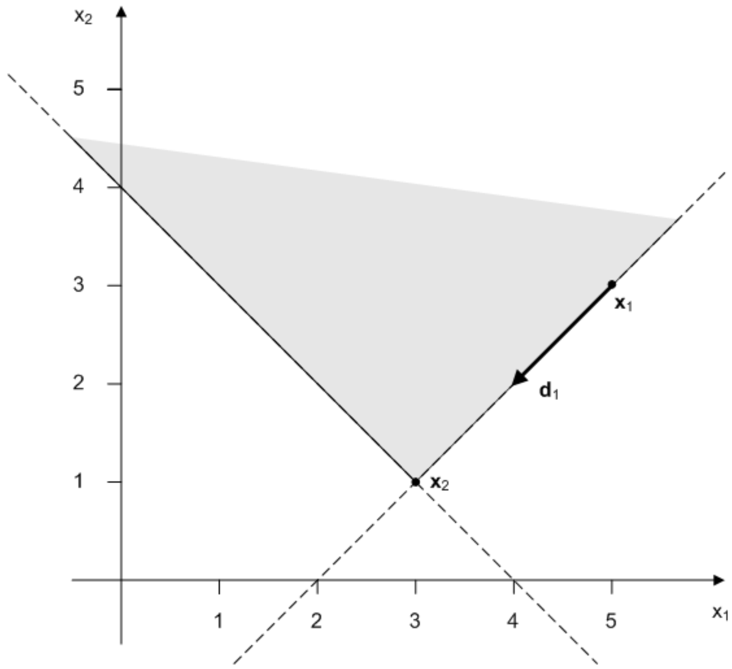
$$-(5 - \lambda) - (3 - \lambda) \leqq -4,$$

amelyből $\lambda \leqq 2$, így $\lambda_{\max} = 2$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (5 - \lambda)^2 + (3 - \lambda)^2 \rightarrow \text{min!} \\ 0 &\leqq \lambda \leqq 2 \end{aligned}$$

Az optimális megoldás $\lambda_1 = 2$, így a következő pont $\mathbf{x}_2 = \mathbf{x}_1 + \lambda_1 \mathbf{d}_1 = (3, 1)$.

Az első lépés eredményeit mutatja az alábbi ábra:



$k = 2$

Az $\mathbf{x}_2 = (3,1)$ lehetséges megoldásnál az első és a második feltétel is aktív. A célfüggvény gradiense az $\mathbf{x}_2 = (3,1)$ pontban $\nabla f(\mathbf{x}_2) = (6,2)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra, ill. a $\bar{\mathbf{d}}$ vektorra a következő:

$$\begin{array}{c} 6d_1 + 2d_2 \to \min! \\ d_1 - d_2 \leqq 0 \\ -d_1 - d_2 \leqq 0 \\ -1 \leqq d_1 \leqq 1 \\ -1 \leqq d_2 \leqq 1 \end{array} \qquad \begin{array}{c} 6\bar{d}_1 + 2\bar{d}_2 - 8 \to \min! \\ \bar{d}_1 - \bar{d}_2 \leqq 0 \\ \bar{d}_1 + \bar{d}_2 \geqq 2 \\ \bar{d}_1 \leqq 2 \\ \bar{d}_2 \leqq 2 \\ \bar{d}_1, \bar{d}_2 \geqq 0 \end{array}$$

A feladatot szimplex módszerrel oldjuk meg. Az induló és a további szimplex tábla:

	$\bar{d}_1$	$\bar{d}_2$	v	
u_1	1	-1	0	0
u_2^*	1	$\boxed{1}$	-1	2
u_3	1	0	0	2
u_4	0	1	0	2
	-6	-2	0	-8
	1	1	-1	2

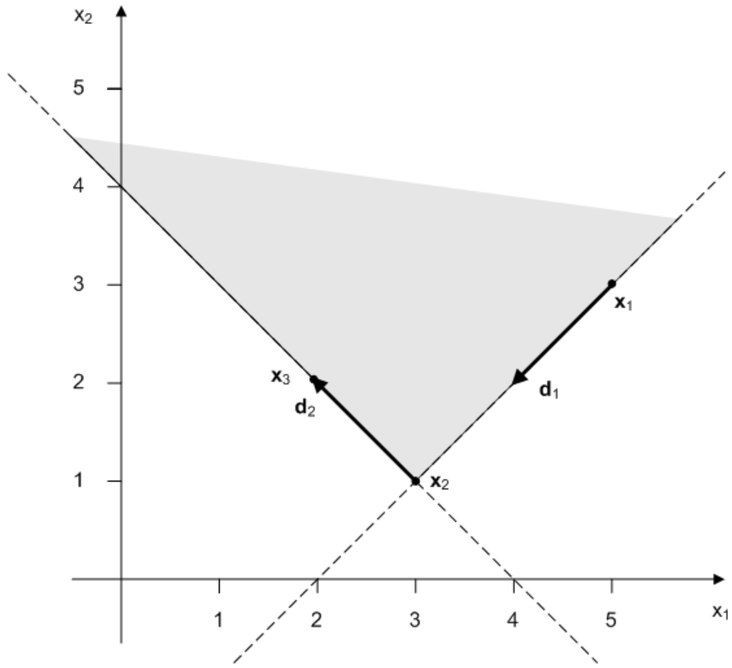
	$\bar{d}_1$	v	
u_1	2	-1	2
$\bar{d}_2$	1	-1	2
u_3	1	0	2
u_4	-1	1	0
	-4	-2	-4

A fenti lineáris programozási feladat optimális megoldása: $\bar{d}_1 = 0$, $\bar{d}_2 = 2$. A következő irányvektor vektorosan: $\mathbf{d}_2 = (-1, 1)$. A célfüggvény minimális értéke: -4, mivel nem zérus, így az $\mathbf{x}_1 = (1, 3)$ megoldás nem KKT pont.

A $\mathbf{d}_2 = (-1, 1)$ irányban az új $\mathbf{x}$ vektor $\mathbf{x} = \mathbf{x}_2 + \lambda \mathbf{d}_2 = (3 - \lambda, 1 + \lambda)$. A $\lambda_{\max} = +\infty$, mert nincs inaktív feltétel. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (3-\lambda)^2 + (1+\lambda)^2 \to \min! \\ \lambda &\geqq 0 \end{aligned}$$

Ennek optimális megoldása $\lambda_2 = 1$, így a következő pont $\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2 \mathbf{d}_2 = (2, 2)$. A második lépés eredményeit mutatja az alábbi ábra:



$k = 3$

Az $\mathbf{x}_3 = (2, 2)$ lehetséges megoldásnál az első feltétel inaktív, a második feltétel aktív. A célfüggvény gradiense az $\mathbf{x}_3 = (2, 2)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_3) = (4, 4)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra, ill. a $\bar{\mathbf{d}}$ vektorra a következő:

$$\begin{array}{c} 4d_1 + 4d_2 \to \min! \\ -d_1 - d_2 \leqq 0 \\ -1 \leqq d_1 \leqq 1 \\ -1 \leqq d_2 \leqq 1 \end{array} \qquad \begin{array}{c} 4\bar{d}_1 + 4\bar{d}_2 - 8 \to \min! \\ \bar{d}_1 + \bar{d}_2 \geqq 2 \\ \bar{d}_1 \leqq 2 \\ \bar{d}_2 \leqq 2 \\ \bar{d}_1, \bar{d}_2 \geqq 0 \end{array}$$

A feladatot szimplex módszerrel oldjuk meg. Az induló és a további szimplex tábla:

	$\bar{d}_1$	$\bar{d}_2$	v	
u_1^*	$\boxed{1}$	1	-1	2
u_2	1	0	0	2
u_2	0	1	0	2
	-4	-4	0	-8
	1	1	-1	2

	$\bar{d}_2$	v	
$\bar{d}_1$	1	-1	2
u_2	-1	1	0
u_2	1	0	2
	0	-4	0

A fenti lineáris programozási feladat optimális megoldása: $\bar{d}_1 = 2,\ \bar{d}_2 = 0$. A vizsgálósorban a $\bar{d}_2$ változó alatt zérus áll, ez azt jelenti, hogy alternatív optimumok vannak. Ha ebben az oszlopban választunk pivotelemet, akkor az optimális célfüggvényérték nem változik, az optimális megoldás pedig: $\bar{d}_1 = 0,\ \bar{d}_2 = 2$. Azt is tudjuk, hogy alternatív optimumok esetén az optimális megoldások konvex lineáris kombinációja is optimális megoldás. Válasszuk a két megoldás számtani átlagát, ami egy konvex kombináció, így $\bar{d}_1 = 1,\ \bar{d}_2 = 1$. A következő irányvektor vektorosan: $\mathbf{d}_3 = (0,0)$. A minimális célfüggvényérték $\min(\boldsymbol{\nabla} f(\mathbf{x}_3)\mathbf{d}_3) = 0$. Megállunk, mivel az optimális célfüggvényérték zérus, így a tétel értelmében az $\mathbf{x}_3 = (2,2)$ lehetséges megoldás KKT pont.

3.2.2. Nemlineáris feltételek esete

Ebben a részben olyan optimalizálási feladattal foglalkozunk, amelyben a feltételek nem szükségképpen lineárisak és ezek egyenlőtlenségek. Az egyenlőséges feltételek kezelése nemlineáris esetben bonyolultabb, ezért e helyen nem foglalkozunk velük.

A feltételes optimalizálási feladat alakja a szokásos formában:

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \min! \\ g_i(\mathbf{x}) &\leqq 0, \quad i = 1,\ldots,m \end{aligned}$$

TÉTEL

Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennáll, hogy $g_i(\mathbf{x}) \leqq 0,\ i = 1,\ldots,m$ és legyen az aktív feltételek indexhalmaza $I = \{i : g_i(\mathbf{x}) = 0\}$. Továbbá tegyük fel, hogy $f, g_i\ i \in I$ függvények differenciálhatók az $\mathbf{x}$ pontban és a $g_i\ i \notin I$ függvények folytonosak az $\mathbf{x}$ pontban. Ha

$$\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} < 0,\ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d} < 0,\ i \in I,$$

akkor a $\mathbf{d}$ vektor javító lehetséges irány.

Bizonyítás

A javító és a lehetséges irányok karakterizációs tételeinek összevonásából azonnal adódik a tétel. A karakterizációs tételek bizonyítása egy másik tananyagban megtalálhatók. **Q.e.d.**

A tétel szerint tehát egy javító lehetséges irányra fenn kell állnia, hogy

$$\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} < 0,\ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d} < 0,\ i \in I.$$

A fenti feltételeket kielégítő irány keresésére természetes módon adódik az alábbi optimalizálási feladat

$$\max\{\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d},\ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d},\ i \in I\} \rightarrow \min!$$

Tehát keressük azt a $\mathbf{d}$ irányvektort, amely a $\boldsymbol{\nabla} f(\mathbf{x})\mathbf{d},\ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d},\ i \in I$ iránymenti deriváltak legnagyobbika is minél kisebb. Ez a feladat kezelhető formába egy z változó bevezetésével tehető, legyen z az iránymenti deriváltak legnagyobbika, ekkor az iránykeresésre az alábbi lineáris programozási feladat (LP) szolgál:

$$\begin{aligned} z &\rightarrow \min! \\ \boldsymbol{\nabla} f(\mathbf{x})\mathbf{d} &\leqq z, \\ \boldsymbol{\nabla} g_i(\mathbf{x})\mathbf{d} &\leqq z,\ i \in I \end{aligned}$$

E feladatnál is szükséges a $\mathbf{d}$ vektorra valamilyen normalizáló feltételt tenni. Legyen ez a koordináták $[-1,1]$ intervallumban tartása. Az előző példákból tapasztaltuk, hogy ezt tudjuk legkényelmesebben kezelni. Ekkor az alábbi lineáris programozási feladat (LP) megoldása szolgáltatja a javító lehetséges irányt:

$$\begin{aligned} z &\rightarrow \min! \\ \nabla f(\mathbf{x})\mathbf{d} - z &\leqq 0, \\ \nabla g_i(\mathbf{x})\mathbf{d} - z &\leqq 0,\ i \in I \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \ldots, n \end{aligned}$$

TÉTEL

Tekintsük az alábbi feltételes optimalizálási feladatot

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \min! \\ g_i(\mathbf{x}) &\leqq 0, \quad i = 1, \ldots, m \end{aligned}$$

Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennáll, hogy $g_i(\mathbf{x}) \leqq 0,\ i = 1, \ldots, m$ és legyen az aktív feltételek indexhalmaza $I = \{i : g_i(\mathbf{x}) = 0\}$. Ekkor az $\mathbf{x}$ lehetséges megoldás akkor és csak akkor Fritz John (FJ) pont, ha a

$$\begin{aligned} z &\rightarrow \min! \\ \nabla f(\mathbf{x})\mathbf{d} - z &\leqq 0, \\ \nabla g_i(\mathbf{x})\mathbf{d} - z &\leqq 0,\ i \in I \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \ldots, n \end{aligned}$$

lineáris programozási feladat célfüggvényének optimális (minimális) értéke zérus.

Bizonyítás

Az iránykereső feladat célfüggvényének optimális értéke akkor és csak akkor zérus, ha a

$$\nabla f(\mathbf{x})\mathbf{d} < 0,\ \nabla g_i(\mathbf{x})\mathbf{d} < 0,\ i \in I$$

rendszernek nincs megoldása. Ekkor pedig a Gordan tétel szerint léteznek olyan $u_0, \mu_i\ i \in I$ számok, hogy

$$\begin{aligned} \mu_0 \nabla f(\mathbf{x}) + \sum_{i \in I} \mu_i \nabla g_i(\mathbf{x}) &= \mathbf{0} \\ \mu_0, \mu_i &\geqq 0\ \ i \in I \\ (\mu_0, \mathbf{u}_I) &\neq \mathbf{0} \end{aligned}$$

ahol $\mathbf{u}_I$ egy olyan vektor, amelynek komponensei $\mu_i \geqq 0,\ i \in I$. A fenti összefüggések pedig pontosan a FJ feltételeket adják.

Zoutendijk módszer algoritmusa nemlineáris feltételek esetén:

Induló lépés ($k = 1$): Kiindulunk egy $\mathbf{x}_k$ lehetséges megoldásból, amelyre $g_i(\mathbf{x}_k) \leqq 0,\ \ i = 1, \ldots, m$.

Közbülső lépés:

1. Legyen az aktív feltételek indexhalmaza $I = \{i : g_i(\mathbf{x}_k) = 0, \quad i = 1, \dots, m\}$.

   (a) Ha $I = \emptyset$, akkor minden $\mathbb{R}^n$-beli vektor lehetséges irány, válasszuk a legjobb javító irányt, azaz legyen $\mathbf{d}_k = -\boldsymbol{\nabla} f(\mathbf{x}_k)$.

   (b) Ha $I \neq \emptyset$, akkor meghatározzuk az $\mathbf{x}_k$ lehetséges megoldáshoz a javító lehetséges irányt, azaz megoldjuk az alábbi lineáris programozási feladatot:

$$\begin{aligned} z &\rightarrow \text{min!} \\ \boldsymbol{\nabla} f(\mathbf{x}_k)\mathbf{d} - z &\leqq 0, \\ \boldsymbol{\nabla} g(\mathbf{x}_k)\mathbf{d} - z &\leqq 0, \quad i \in I, \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \dots, n \end{aligned}$$

   Legyen az optimális megoldás $(z_k, \mathbf{d}_k)$.

2. Megállunk, ha az optimális célfüggvény érték $z_k = 0$. Ekkor az $\mathbf{x}_k$ lehetséges megoldás FJ pont. Egyébként folytatjuk az eljárást a következő sorral.

3. Megoldjuk a $\mathbf{d}_k$ javító lehetséges irányban az iránymenti optimalizálási feladatot, amely az alábbi:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda \mathbf{d}_k) \rightarrow \text{min!} \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

   ahol

$$\lambda_{\max} = \sup \{\lambda : g_i(\mathbf{x}_k + \lambda \mathbf{d}_k) \leqq 0, \quad i = 1, \dots, m\}.$$

   Legyen az optimális megoldás λ_k.

4. Meghatározzuk a következő közelítést: $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k$.

5. Folytatjuk az eljárást, $k := k + 1$.

8. Példa

Legyenek adott az $x_2 = x_1^2$ parabola és az origó centrumú $\sqrt{20}$ sugarú kör. Tekintsük azt a tartományt, amely a parabolán belüli és a körtartomány közös része. Határozzuk meg a tartomány azon pontját, amely az $(5, 4)$ ponthoz a legközelebb van! Oldjuk meg az optimalizálási feladatot Zoutendijk módszerrel!

Megoldás:

Először felírjuk a feladatot matematikai formában. A parabolatartomány: $x_2 \geqq x_1^2$, a körtartomány: $x_1^2 + x_2^2 \leqq 20$. A távolság helyett annak négyzetét használhatjuk, így a célfüggvény $(x_1 - 5)^2 + (x_2 - 4)^2$. A megoldandó optimalizálási feladat tehát a következő:

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 - 10x_1 - 8x_2 \rightarrow \text{min!} \\ g_1(x_1, x_2) &= x_1^2 - x_2 \leqq 0 \\ g_2(x_1, x_2) &= x_1^2 + x_2^2 - 20 \leqq 0 \end{aligned}$$

Megoldás

A célfüggvény és a feltételi függvények gradiens vektora:

$$\begin{aligned}\nabla f &= (2x_1 - 10, 2x_2 - 8)\\ \nabla g_1 &= (2x_1, -1)\\ \nabla g_2 &= (2x_1, 2x_2)\end{aligned}$$

$k = 1$

Legyen az $\mathbf{x}_1 = (1, 1)$ egy lehetséges megoldás. Az első feltétel aktív, a második feltétel inaktív. A célfüggvény és az aktív feltételi függvény gradiense az $\mathbf{x}_1 = (1, 1)$ pontban $\nabla f(\mathbf{x}_1) = (-8, -6), \nabla g_1(\mathbf{x}_1) = (2, -1)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő (az irányvektor koordinátáira tettünk normáló feltételt):

$$\begin{aligned} z &\rightarrow \min!\\ -8d_1 - 6d_2 - z &\leqq 0\\ 2d_1 - d_2 - z &\leqq 0\\ -1 &\leqq d_1 \leqq 1\\ -1 &\leqq d_2 \leqq 1\end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $d_1 = -1/2, d_2 = 1, z = -2$. Mivel z nem zérus, így az $\mathbf{x}_1 = (1, 1)$ megoldás nem FJ pont. Az $\mathbf{x}_1 = (1, 1)$ ponthoz tartozó $\mathbf{d}_1$ irányvektor vektorosan és a hosszát megváltoztatva: $\mathbf{d}_1 = (-1, 2)$. A $\mathbf{d}_1 = (-1, 2)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_1 + \lambda\mathbf{d}_1 = (1-\lambda, 1+2\lambda)$. A $\lambda_{\max}$ értékének meghatározását úgy végezzük, hogy az összes feltételbe behelyettesítjük az $\mathbf{x} = (1 - \lambda, 1 + 2\lambda)$ vektort és meghatározzuk azt a λ-t, amelynél még érvényesek a feltételek, jelen példában

$$\begin{aligned}(1-\lambda)^2 - (1+2\lambda) &\leqq 0,\\ (1-\lambda)^2 + (1+2\lambda)^2 - 20 &\leqq 0,\end{aligned}$$

amelyből az első feltételnél $0 \leqq \lambda \leqq 4$, a második feltételnél $-2.1 \leqq \lambda \leqq 1.7$, amelyből $\lambda_{\max} = 1.7$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned}\varphi(\lambda) &= (1-\lambda)^2 + (1+2\lambda)^2 - 10(1-\lambda) - 8(1+2\lambda) \rightarrow \min!\\ 0 &\leqq \lambda \leqq 1.7\end{aligned}$$

Az optimális megoldás $\lambda_1 = 0.4$, így a következő pont $\mathbf{x}_2 = \mathbf{x}_1 + \lambda_1\mathbf{d}_1 = (0.6, 1.8)$.

$k = 2$

Az $\mathbf{x}_2 = (0.6, 1.8)$ lehetséges megoldásnál mindkét feltétel inaktív, tehát az $\mathbf{x}_2$ pont belső pont, így minden irány megengedett, célszerű a legjobb javító irányt választani, amely a célfüggvény gradiensének (-1)-szerese. A célfüggvény gradiense az $\mathbf{x}_2 = (0.6, 1.8)$ pontban $\nabla f(\mathbf{x}_2) = (-8.8, -4.4)$. A $\mathbf{d}_2$ javító irány, egyszerűbb formában írva, $\mathbf{d}_2 = (2, 1)$. A $\mathbf{d}_2 = (2, 1)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_2 + \lambda\mathbf{d}_2 = (0.6 + 2\lambda, 1.8 + \lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\begin{aligned}(0.6+2\lambda)^2 - (1.8+\lambda) &\leqq 0,\\ (0.6+2\lambda)^2 + (1.8+\lambda)^2 - 20 &\leqq 0,\end{aligned}$$

amelyből az első feltételnél $-0.8 \leqq \lambda \leqq 0.45$, a második feltételnél $-2.5 \leqq \lambda \leqq 1.3$, amelyből $\lambda_{\max} = 0.45$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (0.6+2\lambda)^2 + (1.8+\lambda)^2 - 10(0.6+2\lambda) - 8(1.8+\lambda) \rightarrow \min! \\ 0 &\leqq \lambda \leqq 0.45 \end{aligned}$$

Az optimális megoldás $\lambda_2 = 0.45$, így a következő pont $\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2 \mathbf{d}_2 = (1.5, 2.25)$.

$k = 3$

Az $\mathbf{x}_3 = (1.5, 2.25)$ lehetséges megoldásnál az első feltétel aktív, a második feltétel inaktív. A célfüggvény és az aktív feltételi függvény gradiense az $\mathbf{x}_3 = (1.5, 2.25)$ pontban $\nabla f(\mathbf{x}_3) = (-7, -3.5), \nabla g_1(\mathbf{x}_3) = (3, -1)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő:

$$\begin{aligned} z &\rightarrow \min! \\ -7d_1 - 3.5d_2 - z &\leqq 0 \\ 3d_1 - d_2 - z &\leqq 0 \\ -1 &\leqq d_1 \leqq 1 \\ -1 &\leqq d_2 \leqq 1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $\mathbf{d}_3 = (-0.25, 1), z = -1.75$. Mivel z nem zérus, így az $\mathbf{x}_3 = (1.5, 2.25)$ megoldás nem FJ pont. Az egyszerübb számolás miatt a $\mathbf{d}_3 = (-1, 4)$ irányvektort használjuk. A $\mathbf{d}_3 = (-1, 4)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_3 + \lambda \mathbf{d}_3 = (1.5 - \lambda, 2.25 + 4\lambda)$. A $\lambda_{\max}$ értékének meghatározása:

$$\begin{aligned} (1.5-\lambda)^2 - (2.25+4\lambda) &\leqq 0, \\ (1.5-\lambda)^2 + (2.25+4\lambda)^2 - 20 &\leqq 0, \end{aligned}$$

amelyből az első feltételnél $0 \leqq \lambda \leqq 7$, a második feltételnél $-1.4 \leqq \lambda \leqq 0.53$, amelyből $\lambda_{\max} = 0.53$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (1.5-\lambda)^2 + (2.25+4\lambda)^2 - 10(1.5-\lambda) - 8(2.25+4\lambda) \rightarrow \min! \\ 0 &\leqq \lambda \leqq 0.53 \end{aligned}$$

Az optimális megoldás $\lambda_3 = 0.2$, így a következő pont $\mathbf{x}_4 = \mathbf{x}_3 + \lambda_3 \mathbf{d}_3 = (1.3, 3.05)$.

$k = 4$

Itt abbahagyjuk az algoritmust, addig kellene folytatni az eljárást, amíg a lineáris programozási feladat optimális célfüggvényének értékére $z = 0$ adódik. Ekkor a tétel értelmében az aktuális $\mathbf{x}_k$ megoldás Fritz John pont.

3.2.3. Topkis-Veinott módszer

A Zoutendijk algoritmus nemlineáris feltételek esetén nem minden esetben konvergál a Fritz-John ponthoz. Topkis és Veinott módosítottak az algoritmuson úgy, hogy a nem aktív feltételeket is figyelembe veszik az irány meghatározásánál, ezzel garantálják a konvergenciát.

Topkis-Veinott módszer algoritmusa:

Induló lépés ($k = 1$): Kiindulunk egy $\mathbf{x}_k$ lehetséges megoldásból, amelyre $g_i(\mathbf{x}_k) \leqq 0, \quad i = 1, \ldots, m$.

Közbülső lépés:

1. Legyen az aktív feltételek indexhalmaza $I = \{i : g_i(\mathbf{x}_k) = 0, \quad i = 1, \ldots, m\}$.
   (a) Ha $I = \emptyset$, akkor minden $\mathbb{R}^n$-beli vektor lehetséges irány, válasszuk a legjobb javító irányt, azaz legyen $\mathbf{d}_k = -\boldsymbol{\nabla} f(\mathbf{x}_k)$.
   (b) Ha $I \neq \emptyset$, akkor meghatározzuk az $\mathbf{x}_k$ lehetséges megoldáshoz a javító lehetséges irányt, azaz megoldjuk az alábbi lineáris programozási feladatot:

$$\begin{aligned} z &\rightarrow \text{min!} \\ \boldsymbol{\nabla} f(\mathbf{x}_k)\mathbf{d} - z &\leqq 0, \\ \boldsymbol{\nabla} g(\mathbf{x}_k)\mathbf{d} - z &\leqq -g(\mathbf{x}_k), \quad i = 1, \ldots, m, \\ -1 &\leqq d_i \leqq 1, \quad i = 1, \ldots, n \end{aligned}$$

   Legyen az optimális megoldás $(z_k, \mathbf{d}_k)$.

2. Megállunk, ha az optimális célfüggvény érték $z_k = 0$. Ekkor az $\mathbf{x}_k$ lehetséges megoldás FJ pont. Egyébként folytatjuk az eljárást a következő sorral.

3. Megoldjuk a $\mathbf{d}_k$ javító lehetséges irányban az iránymenti optimalizálási feladatot, amely az alábbi:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda \mathbf{d}_k) \rightarrow \text{min!} \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

   ahol

$$\lambda_{\max} = \sup \{\lambda : g_i(\mathbf{x}_k + \lambda \mathbf{d}_k) \leqq 0, \quad i = 1, \ldots, m\}.$$

   Legyen az optimális megoldás λ_k.

4. Meghatározzuk a következő közelítést: $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k$.

5. Folytatjuk az eljárást, $k := k + 1$.

TÉTEL
Legyen az f, g_i $i = 1, \ldots, m$ függvények folytonosan differenciálhatók. Ekkor a Topkis és Veinott algoritmus által generált $\{\mathbf{x}_k\}$ sorozat minden torlódási pontja Fritz-John pont.

A következőkben az előző feladatot Topkis-Veinott algoritmussal is megoldjuk.

9. Példa
Oldjuk meg a 8. példabeli optimalizálási feladatot Topkis-Veinott módszerrel!

$$\begin{aligned} f(x_1, x_2) &= x_1^2 + x_2^2 - 10x_1 - 8x_2 \rightarrow \text{min!} \\ g_1(x_1, x_2) &= x_1^2 - x_2 \leqq 0 \\ g_2(x_1, x_2) &= x_1^2 + x_2^2 - 20 \leqq 0 \end{aligned}$$

Megoldás
A célfüggvény és a feltételi függvények gradiens vektora:

$$\begin{aligned} \boldsymbol{\nabla} f &= (2x_1 - 10, 2x_2 - 8) \\ \boldsymbol{\nabla} g_1 &= (2x_1, -1) \\ \boldsymbol{\nabla} g_2 &= (2x_1, 2x_2) \end{aligned}$$

$k = 1$

Legyen az $\mathbf{x}_1 = (1,1)$ egy lehetséges megoldás. Az első feltétel aktív, a második feltétel inaktív, itt $g_2(\mathbf{x}_1) = -18$. A célfüggvény és az összes feltételi függvény gradiense az $\mathbf{x}_1 = (1,1)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_1) = (-8,-6), \boldsymbol{\nabla} g_1(\mathbf{x}_1) = (2,-1), \boldsymbol{\nabla} g_2(\mathbf{x}_1) = (2,2)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő (az irányvektor koordinátáira tettünk normáló feltételt):

$$\begin{aligned} z &\rightarrow \min! \\ -8d_1 - 6d_2 - z &\leqq 0 \\ 2d_1 - d_2 - z &\leqq 0 \\ 2d_1 + 2d_2 - z &\leqq 18 \\ -1 &\leqq d_1 \leqq 1 \\ -1 &\leqq d_2 \leqq 1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $d_1 = -1/2, d_2 = 1, z = -2$. Mivel z nem zérus, így az $\mathbf{x}_1 = (1,1)$ megoldás nem FJ pont. Az $\mathbf{x}_1 = (1,1)$ ponthoz tartozó $\mathbf{d}_1$ irányvektor vektorosan és a hosszát megváltoztatva: $\mathbf{d}_1 = (-1,2)$. A $\mathbf{d}_1 = (-1,2)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_1 + \lambda \mathbf{d}_1 = (1-\lambda, 1+2\lambda)$. A $\lambda_{\max}$ értékének meghatározása hasonlóan történik, ahogy azt az előző példánál láttuk, így $\lambda_{\max} = 1.7$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (1-\lambda)^2 + (1+2\lambda)^2 - 10(1-\lambda) - 8(1+2\lambda) \rightarrow \min! \\ 0 &\leqq \lambda \leqq 1.7 \end{aligned}$$

Az optimális megoldás $\lambda_1 = 0.4$, így a következő pont $\mathbf{x}_2 = \mathbf{x}_1 + \lambda_1 \mathbf{d}_1 = (0.6, 1.8)$.

$k = 2$

Az $\mathbf{x}_2 = (0.6, 1.8)$ lehetséges megoldásnál mindkét feltétel inaktív, $g_1(\mathbf{x}_2) = -1.44$, $g_2(\mathbf{x}_2) = -16.4$. A célfüggvény és az összes feltételi függvény gradiense az $\mathbf{x}_2 = (0.6, 1.8)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_1) = (-8.8, -4.4), \boldsymbol{\nabla} g_1(\mathbf{x}_1) = (1.2, -1), \boldsymbol{\nabla} g_2(\mathbf{x}_1) = (1.2, 3.6)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő:

$$\begin{aligned} z &\rightarrow \min! \\ -8.8d_1 - 4.4d_2 - z &\leqq 0 \\ 1.2d_1 - d_2 - z &\leqq 1.44 \\ 1.2d_1 + 3.6d_2 - z &\leqq 16.4 \\ -1 &\leqq d_1 \leqq 1 \\ -1 &\leqq d_2 \leqq 1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $d_1 = -0.2,\ d_2 = 1, z = -2.6$. Mivel z nem zérus, így az $\mathbf{x}_2 = (0.6, 1.8)$ megoldás nem FJ pont. Az $\mathbf{x}_2 = (0.6, 1.8)$ ponthoz tartozó $\mathbf{d}_2$ irányvektor $\mathbf{d}_2 = (-1, 5)$. A $\mathbf{d}_2 = (-1, 5)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_2 + \lambda \mathbf{d}_2 = ((0.6 - \lambda, 1.8 + 5\lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\begin{aligned} (0.6-\lambda)^2 - (1.8+5\lambda) &\leqq 0, \\ (0.6-\lambda)^2 + (1.8+5\lambda)^2 - 20 &\leqq 0, \end{aligned}$$

amelyből az első feltételnél $-0.2 \leqq \lambda \leqq 6.4$, a második feltételnél $-1.2 \leqq \lambda \leqq 0.53$, amelyből $\lambda_{\max} = 0.53$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (0.6-\lambda)^2 + (1.8+5\lambda)^2 - 10(0.6-\lambda) - 8(1.8+5\lambda) \to \min! \\ 0 &\leqq \lambda \leqq 0.53 \end{aligned}$$

Az optimális megoldás $\lambda_2 = 0.25$, így a következő pont $\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2 \mathbf{d}_2 = (0.35, 3.05)$.

$k = 3$

Az $\mathbf{x}_3 = (0.35, 3.05)$ lehetséges megoldásnál az első és a második feltétel is inaktív, itt $g_1(\mathbf{x}_3) = -2.9$, $g_2(\mathbf{x}_3) = -10.6$. A célfüggvény és az összes feltételi függvény gradiense az $\mathbf{x}_3 = (0.35, 3.05)$ pontban $\nabla f(\mathbf{x}_3) = (-9.3, -1.9)$, $\nabla g_1(\mathbf{x}_3) = (0.7, -1)$, $\nabla g_2(\mathbf{x}_3) = (0.7, 6.1)$. A megoldandó lineáris programozási feladat a $\mathbf{d} = (d_1, d_2)$ irányvektorra a következő:

$$\begin{aligned} z &\to \min! \\ -9.3d_1 - 1.9d_2 - z &\leqq 0 \\ 0.7d_1 - d_2 - z &\leqq 2.9 \\ 0.7d_1 + 6.1d_2 - z &\leqq 10.6 \\ -1 &\leqq d_1 \leqq 1 \\ -1 &\leqq d_2 \leqq 1 \end{aligned}$$

A fenti lineáris programozási feladat optimális megoldása: $d_1 = 0.2$, $d_2 = 1, z = -3.76$. Mivel z nem zérus, így az $\mathbf{x}_3 = (0.35, 3.05)$ megoldás nem FJ pont. Az $\mathbf{x}_3 = (0.35, 3.05)$ ponthoz tartozó $\mathbf{d}_3$ irányvektor: $\mathbf{d}_3 = (1, 5)$. A $\mathbf{d}_3 = (1, 5)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_3 + \lambda \mathbf{d}_3 = (0.35 + \lambda, 3.05 + 5\lambda)$. A feltélelekbe behelyettesítve kapjuk, hogy $\lambda_{\max} = 0.275$. Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (0.35+\lambda)^2 + (3.05+5\lambda)^2 - 10(0.35+\lambda) - 8(3.05+5\lambda) \to \min! \\ 0 &\leqq \lambda \leqq 0.275 \end{aligned}$$

Az optimális megoldás $\lambda_3 = 0.275$, így a következő pont $\mathbf{x}_4 = \mathbf{x}_3 + \lambda_3 \mathbf{d}_3 = (0.625, 4.425)$.

$k = 4$

Itt abbahagyjuk az algoritmust, addig kellene folytatni, amíg a lineáris programozási feladat optimális célfüggvényének értékére $z = 0$ adódik. A tétel értelmében az algoritmus konvergens és Fritz John ponthoz konvergál.

Feladat:

Rajzolja fel a nemlineáris feltételes feladat megengedett tartományát!

A Zoutendijk és a Topkis-Veinott algoritmus egyes lépéseiben kapott eredményeket rajzolja be az ábrába és hasonlítsa össze a két módszert!

3.3. Rosen-féle gradiens projekciós módszer

3.3.1. Projekciós mátrix definíciója

Mielőtt a módszer ismertetésére rátérnénk, definiáljuk az eljárásban használatos projekciós mátrix fogalmát és egy kapcsolódó tételt.

Legyen $\mathbf{P} \in \mathbb{R}^{n\times n}$ négyzetes mátrix. A $\mathbf{P}$ mátrix projekciós, ha

1. $\mathbf{P} = \mathbf{P}^T$ (szimmetrikus),
2. $\mathbf{PP} = \mathbf{P}$ (idempotens).

TÉTEL
a) Ha $\mathbf{P}$ projekciós mátrix, akkor $\mathbf{P}$ pozitív szemidefinit.
b) A $\mathbf{P}$ mátrix akkor és csak akkor projekciós, ha $\mathbf{E}-\mathbf{P}$ projekciós mátrix, ahol $\mathbf{E} \in \mathbb{R}^{n\times n}$ egységmátrix.

Bizonyítás

a) $\mathbf{x}^T\mathbf{Px} = \mathbf{x}^T\mathbf{PPx} = \mathbf{x}^T\mathbf{P}^T\mathbf{Px} = (\mathbf{Px})^T\mathbf{Px} = (\mathbf{Px})^2 \geqq 0$. Tehát, ha $\mathbf{P}$ projekciós, akkor $\mathbf{x}^T\mathbf{Px} \geqq 0$ minden $\mathbf{x}$, így minden $\mathbf{x} \neq \mathbf{0}$ vektorra is, vagyis $\mathbf{P}$ pozitív szemidefinit.

b) Ha $\mathbf{P}$ projekciós, akkor $\mathbf{E}-\mathbf{P} = \mathbf{E}^T - \mathbf{P}^T = (\mathbf{E}-\mathbf{P})^T$ és $(\mathbf{E}-\mathbf{P})(\mathbf{E}-\mathbf{P}) = \mathbf{E}-\mathbf{P}-\mathbf{P}+\mathbf{PP} = \mathbf{E}-\mathbf{P}-\mathbf{P}+\mathbf{P} = \mathbf{E}-\mathbf{P}$. Hasonlóan bizonyítható a másik irány is. **Q.e.d.**

Először olyan feltételes optimalizálási feladatokat vizsgálunk, amelyekben a feltételi függvények lineárisak.

Mint majd látni fogjuk a javító lehetséges irányt a célfüggvény gradiens vektorának a vetítése adja, ezért szerepelnek a módszer nevében a gradiens és a projekció (vetítés) szavak.

3.3.2. Lineáris feltételek esete

A feltételes optimalizálási feladat alakja a szokásos formában

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \min! \\ \mathbf{Ax} &\leqq \mathbf{b} \\ \mathbf{Bx} &= \mathbf{c} \end{aligned}$$

ahol $f : \mathbb{R}^n \rightarrow \mathbb{R}, \mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{B} \in \mathbb{R}^{k\times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^k$.

Legyen adott egy $\mathbf{x}$ lehetséges megoldás, ehhez a legjobb javító irány a $-\nabla f(\mathbf{x})$, de ez nem feltétlenül lehetséges irány. Az alábbi tétel állítása szerint, ha a legjobb javító irányt egy alkalmas $\mathbf{P}$ projekciós mátrix-szal megszorozzuk, akkor az így nyert $\mathbf{d} = -\mathbf{P}\nabla f(\mathbf{x})$ vektor javító és lehetséges irány lesz.

TÉTEL
Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennáll, hogy $\mathbf{Ax} \leqq \mathbf{b}, \mathbf{Bx} = \mathbf{c}$. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x} < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói.

a) Ha $\mathbf{P}$ projekciós mátrix és $\mathbf{P}\nabla f(\mathbf{x}) \neq \mathbf{0}$, akkor $\mathbf{d} = -\mathbf{P}\nabla f(\mathbf{x})$ javító irány.

b) Legyen $\mathbf{M} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{B} \end{bmatrix}$ és sorvektorai legyenek lineárisan függetlenek, legyen továbbá $\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{MM}^T)^{-1}\mathbf{M}$, ekkor a $\mathbf{d} = -\mathbf{P}\nabla f(\mathbf{x})$ javító lehetséges irány.

Bizonyítás

1. Először megmutatjuk, hogy bármely $\mathbf{P}$ projekciós mátrix esetén $\mathbf{d} = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x})$ javító irány. Azt kell belátni, hogy $\boldsymbol{\nabla} f(\mathbf{x})^T\mathbf{d} < 0$.
$\boldsymbol{\nabla} f(\mathbf{x})^T\mathbf{d} = -\boldsymbol{\nabla} f(\mathbf{x})^T\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = -\boldsymbol{\nabla} f(\mathbf{x})^T\mathbf{P}\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = -\boldsymbol{\nabla} f(\mathbf{x})^T\mathbf{P}^T\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) =$
$= -(\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}))^T\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = -(\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}))^2 < 0$, mivel $\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) \neq 0$, így $\mathbf{d}$ valóban javító irány.

2. Másodszor megmutatjuk, hogy $\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}$ projekciós mátrix. Az előző tétel szerint elegendő megmutatni, hogy $\mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}$ projekciós mátrix,

   i) $\mathbf{P}^T \stackrel{?}{=} \mathbf{P}: \ \left[\mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\right]^T = \mathbf{M}^T\left[(\mathbf{M}\mathbf{M}^T)^{-1}\right]^T(\mathbf{M}^T)^T = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M},$

   ii) $\mathbf{P}\mathbf{P} \stackrel{?}{=} \mathbf{P}: \ \left[\mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\right]\left[\mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\right] =$
   $= \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}(\mathbf{M}\mathbf{M}^T)(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M} = \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}.$

3. Most pedig megmutatjuk, hogy $\mathbf{M}\mathbf{P} = \mathbf{0}$. $\mathbf{M}\mathbf{P} = \mathbf{M}\left[\mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\right] = \mathbf{M} - \mathbf{M} = \mathbf{0}$. Az $\mathbf{M}\mathbf{P} = \mathbf{0}$ azt fejezi ki, hogy $\mathbf{P}$ mátrix az $\mathbf{M}$ mátrix sorvektorait (pontosabban az $\mathbf{A}_1$ és a $\mathbf{B}$ mátrixok sorvektorait) a zérus vektorba vetíti, más szavakkal a $\mathbf{P}$ mátrix az aktív feltételek gradienseit a zérus vektorba vetíti. Vagy más megfogalmazásban az aktív feltételek gradienseinek vektortere merőleges a $\mathbf{P}$ mátrix oszlopvektor terére.

4. Végül pedig megmutatjuk, hogy $\mathbf{M}\mathbf{d} = \mathbf{0}$. $\mathbf{M}\mathbf{d} = -\mathbf{M}\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = -(\mathbf{M}\mathbf{P})\boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{0}$. Az $\mathbf{M}\mathbf{d} = \mathbf{0}$ részletesen írva: $\mathbf{A}_1\mathbf{d} = \mathbf{0}$ és $\mathbf{B}\mathbf{d} = \mathbf{0}$. Korábbról pedig tudjuk, hogy egy $\mathbf{d}$ vektor akkor lehetséges irány, ha $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$ és $\mathbf{B}\mathbf{d} = \mathbf{0}$, ez pedig esetünkben teljesedik, így $\mathbf{d}$ valóban javító lehetséges irány. A $\mathbf{P}$ mátrix a célfüggvény gradiens vektorának (-1)-szeresét (a legjobb javító irányt), azaz a $-\boldsymbol{\nabla} f(\mathbf{x})$ vektort az aktív feltételek gradienseinek vektorterére merőleges altérre vetíti. **Q.e.d.**

TÉTEL

Legyen $\mathbf{x}$ egy lehetséges megoldás, azaz fennál, hogy $\mathbf{A}\mathbf{x} \leqq \mathbf{b}, \mathbf{B}\mathbf{x} = \mathbf{c}$. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x} = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x} < \mathbf{b}_2$, ahol $\mathbf{A}_1, \mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói. Legyen $\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}$, ahol $\mathbf{M} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{B} \end{bmatrix}$ és $\mathbf{M}$ sorvektorai lineárisan függetlenek.

1. Ha $\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) \neq \mathbf{0}$, akkor $\mathbf{d} = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x})$ javító lehetséges irány.

2. Ha $\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{0}$, akkor legyen a $\mathbf{w} = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\boldsymbol{\nabla} f(\mathbf{x})$ vektor $(\mathbf{u}, \mathbf{v})$ alakban partícionálva, ahol $\mathbf{u} = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{A}_1\boldsymbol{\nabla} f(\mathbf{x})$ és $\mathbf{v} = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{B}\boldsymbol{\nabla} f(\mathbf{x})$.

   (a) Ha $\mathbf{u} \geqq \mathbf{0}$, akkor az $\mathbf{x}$ pont KKT pont.

   (b) Ha $\mathbf{u} \ngeqq \mathbf{0}$, akkor legyen $\hat{\mathbf{M}} = \begin{bmatrix} \hat{\mathbf{A}}_1 \\ \mathbf{B} \end{bmatrix}$, ahol $\hat{\mathbf{A}}_1$ olyan mátrix, amely az $\mathbf{A}_1$ mátrixból úgy keletkezik, hogy egyetlen sort elhagyunk belőle, mégpedig azon sorok közül valamelyiket, ahol u_i negatív. Legyen most $\hat{\mathbf{P}} = \hat{\mathbf{E}} - \hat{\mathbf{M}}^T(\hat{\mathbf{M}}\hat{\mathbf{M}}^T)^{-1}\hat{\mathbf{M}}$ és $\mathbf{d} = -\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x})$. Ekkor a $\mathbf{d}$ vektor javító lehetséges irány.

Bizonyítás

1. Az előző tétel alapján $\mathbf{d}$ javító lehetséges irány.

2. $\mathbf{0} = \mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = \left[\mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\right] \boldsymbol{\nabla} f(\mathbf{x}) = \boldsymbol{\nabla} f(\mathbf{x}) + \mathbf{M}^T\mathbf{w} = \boldsymbol{\nabla} f(\mathbf{x}) + \mathbf{A}_1^T\mathbf{u} + \mathbf{B}^T\mathbf{v}$.

   (a) Ha $\mathbf{u} \geqq \mathbf{0}$, akkor az $\mathbf{x}$ pont KKT pont, hiszen teljesednek a KKT feltételek.

   (b) Ha $\mathbf{u} \ngeqq \mathbf{0}$, ekkor válasszunk ki egy negatív elemet, legyen ez az $u_r < 0$. Az $\hat{\mathbf{M}}$ mátrix az $\mathbf{M}$ mátrixból tehát úgy keletkezik, hogy az r-edik sort elhagyjuk. Legyen $\hat{\mathbf{P}} = \hat{\mathbf{E}} - \hat{\mathbf{M}}^T(\hat{\mathbf{M}}\hat{\mathbf{M}}^T)^{-1}\hat{\mathbf{M}}$ és legyen $\mathbf{d} = -\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x})$. Azt, hogy $\mathbf{d}$ javító lehetséges irány, két lépésben bizonyítjuk.

      i. Először megmutatjuk, hogy $\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x}) \neq \mathbf{0}$, azaz $\mathbf{d}$ javító irány.
      Indirekte tegyük fel, hogy $\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{0}$. Ekkor

$$\begin{aligned} &\text{egyrészt:} \quad \mathbf{0} = \hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x}) = \boldsymbol{\nabla} f(\mathbf{x}) + \hat{\mathbf{M}}^T\hat{\mathbf{w}}, \\ &\text{másrészt:} \quad \mathbf{0} = \mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}) = \boldsymbol{\nabla} f(\mathbf{x}) + \mathbf{M}^T\mathbf{w} = \boldsymbol{\nabla} f(\mathbf{x}) + \hat{\mathbf{M}}^T\bar{\mathbf{w}} + u_r(\mathbf{A}_1)_r \end{aligned}$$

      ahol $\hat{\mathbf{w}}, \bar{\mathbf{w}}$ eggyel kevesebb elemű, mint $\mathbf{w}$, $(\mathbf{A}_1)_r$ pedig az $\mathbf{A}_1$ mátrix r-edik sora oszlopvektorként írva. Vonjuk ki a fenti két egyenletet egymásból

$$\mathbf{0} = \hat{\mathbf{M}}^T(\bar{\mathbf{w}} - \hat{\mathbf{w}}) + u_r(\mathbf{A}_1)_r,$$

      amely nem más, mint az $\mathbf{M}^T$mátrix oszlopvektorainak lineáris kombinációja, más szóval az $\mathbf{M}$ mátrix sorvektorainak lineáris kombinációja. Mivel $u_r \neq 0$ ($u_r < 0$), így $\mathbf{M}$ sorvektorai nem lehetnek lineárisan függetlenek, ami ellentmondás, tehát $\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x}) \neq \mathbf{0}$, amiből a $\mathbf{d} = -\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x})$ az előzőek alapján javító irányt jelent.

      ii. Másodszor pedig megmutatjuk, hogy $\mathbf{d}$ lehetséges irány.
      Egyszerű behelyettesítéssel belátható, hogy $\hat{\mathbf{M}}\hat{\mathbf{P}} = \mathbf{0}$. Ebből következik, hogy $\hat{\mathbf{M}}\mathbf{d} = \mathbf{0}$, azaz $\hat{\mathbf{A}}_1\mathbf{d} = \mathbf{0}$ és $\mathbf{B}\mathbf{d} = \mathbf{0}$. Mivel tudjuk, hogy $\mathbf{A}_1\mathbf{d} \leqq \mathbf{0}$ esetén is lehetséges a $\mathbf{d}$ irány, így már elegendő csupán azt belátni, hogy $(\mathbf{A}_1)_r^T\mathbf{d} \leqq \mathbf{0}$. Térjünk vissza a

$$\mathbf{0} = \boldsymbol{\nabla} f(\mathbf{x}) + \hat{\mathbf{M}}^T\bar{\mathbf{w}} + u_r(\mathbf{A}_1)_r$$

      képlethez, szorozzuk be balról az $(\mathbf{A}_1)_r^T\hat{\mathbf{P}}$ vektorral, ekkor kapjuk, hogy

$$0 = (\mathbf{A}_1)_r^T\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x}) + (\mathbf{A}_1)_r^T\hat{\mathbf{P}}\hat{\mathbf{M}}^T\bar{\mathbf{w}} + u_r(\mathbf{A}_1)_r^T\hat{\mathbf{P}}(\mathbf{A}_1)_r.$$

      Mivel $\hat{\mathbf{P}}\hat{\mathbf{M}}^T = (\hat{\mathbf{M}}\hat{\mathbf{P}})^T = \mathbf{0}$ és $\mathbf{d} = -\hat{\mathbf{P}}\boldsymbol{\nabla} f(\mathbf{x})$, így

$$(\mathbf{A}_1)_r^T\mathbf{d} = u_r(\mathbf{A}_1)_r^T\hat{\mathbf{P}}(\mathbf{A}_1)_r.$$

      A $\hat{\mathbf{P}}$ projekciós mátrix, így $\hat{\mathbf{P}}$ pozitív szemidefinit és $u_r < 0$, ebből következik, hogy $(\mathbf{A}_1)_r^T\mathbf{d} \leqq \mathbf{0}$. **Q.e.d.**

Megjegyzés:

A $\mathbf{d}$ javító lehetséges irányra vonatkozó eredményeket az alábbbi okfejtéssel is megkaphatjuk. Keressük azt az irányt az $\mathbf{x}$ lehetséges pontból, amelynél a célfüggvény a legnagyobbat csökken és a $\mathbf{d}$ irány olyan legyen, hogy minden aktív feltétel gradiensére merőleges, azaz

$\mathbf{A}_1\mathbf{d} = \mathbf{0}$ és $\mathbf{Bd} = \mathbf{0}$, összevonva $\mathbf{Md} = \mathbf{0}$. Használjuk a $\mathbf{d}^2 = 1$ normalizálást, ekkor az alábbi optimalizálási feladatot kell megoldanunk a keresett $\mathbf{d}$ vektorra

$$\begin{aligned} \nabla f(\mathbf{x})\mathbf{d} &\rightarrow \min! \\ \mathbf{Md} &= \mathbf{0}, \\ \mathbf{d}^2 - 1 &= 0. \end{aligned}$$

A feladat optimális megoldására az alábbi szükséges feltételeknek (KKT feltételeknek) kell fennállni

$$\begin{aligned} \nabla f(\mathbf{x}) + \mathbf{M}^T\mathbf{v} + 2\mathbf{d}v_0 &= \mathbf{0}, \\ \mathbf{Md} &= \mathbf{0}, \\ \mathbf{d}^2 - 1 &= 0. \end{aligned}$$

Ennek a rendszernek a megoldását $\mathbf{d}$-re az alábbiak szerint végezhetjük. Szorozzuk be $\mathbf{M}$-el az első egyenletet, ekkor

$$\mathbf{M}\nabla f(\mathbf{x}) + \mathbf{MM}^T\mathbf{v} + 2\mathbf{Md}v_0 = \mathbf{0}$$

Mivel $\mathbf{Md} = \mathbf{0}$, és ha az $\mathbf{M}$ mátrix sorvektorai lineárisan függetlenek, akkor létezik az $\mathbf{MM}^T$ mátrix inverze és ekkor a $\mathbf{v}$ Lagrange szorzók vektora a következőképpen írható

$$\mathbf{v} = -(\mathbf{MM}^T)^{-1}\mathbf{M}\nabla f(\mathbf{x}),$$

amelyet, ha a szükséges feltételek első egyenletébe visszahelyettesítünk kapjuk, hogy

$$\mathbf{d} = -\frac{1}{2v_0}\left[\nabla f(\mathbf{x}) - \mathbf{M}^T(\mathbf{MM}^T)^{-1}\mathbf{M}\nabla f(\mathbf{x})\right] = -\frac{1}{2v_0}\left[\mathbf{E} - \mathbf{M}^T(\mathbf{MM}^T)^{-1}\mathbf{M}\right]\nabla f(\mathbf{x}).$$

A v_0 Lagrange szorzó csupán a normáláshoz kell. Tehát $\mathbf{d}$-re ugyanazt az eredményt kaptuk, azaz

$$\begin{aligned} \mathbf{d} &= -\mathbf{P}\nabla f(\mathbf{x}), \text{ ahol} \\ \mathbf{P} &= \mathbf{E} - \mathbf{M}^T(\mathbf{MM}^T)^{-1}\mathbf{M}. \end{aligned}$$

Iránykeresés

A javító lehetséges irányban egy pont meghatározását a Zoutendijk módszernél elmondottak alapján kell itt is elvégezni. Legyen adott egy $\mathbf{x}_k$ lehetséges megoldás és a hozzátartozó $\mathbf{d}_k$ javító lehetséges irány. Összefoglalva az iránymenti optimalizálási feladat a következő:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda\mathbf{d}_k) \rightarrow \min! \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

ahol

$$\lambda_{\max} = \begin{cases} \min\left\{\frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i} : (\mathbf{A}_2\mathbf{d}_k)_i > 0\right\}, & \text{ha } \mathbf{A}_2\mathbf{d}_k \nleqq \mathbf{0} \\ \infty, & \text{ha } \mathbf{A}_2\mathbf{d}_k \leqq \mathbf{0} \end{cases}$$

Rosen-féle gradiens projekciós módszer algoritmusa lineáris feltételek esetén:

Induló lépés ($k = 1$): Kiindulunk egy $\mathbf{x}_k$ lehetséges megoldásból, amelyre $\mathbf{Ax}_k \leqq \mathbf{b}, \mathbf{Bx}_k = \mathbf{c}$.

Közbülső lépés:

1. Tegyük fel, hogy $\mathbf{A}_1\mathbf{x}_k = \mathbf{b}_1$ és $\mathbf{A}_2\mathbf{x}_k < \mathbf{b}_2$, ahol $\mathbf{A}_1$, $\mathbf{A}_2, \mathbf{b}_1, \mathbf{b}_2$ az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor megfelelő partíciói.

2. Meghatározzuk az aktív feltételeket tartalmazó $\mathbf{M}$ mátrixot, legyen $\mathbf{M} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{B} \end{bmatrix}$.

3. Az $\mathbf{M}$ mátrix vizsgálata:

   (a) Ha nincs $\mathbf{M}$ mátrix (pl. nincs egyenlőséges feltétel és minden egyenlőtlenséges feltétel inaktív), akkor

      i. $\boldsymbol{\nabla} f(\mathbf{x}_k) = \mathbf{0}$ esetben megállunk,

      ii. $\boldsymbol{\nabla} f(\mathbf{x}_k) \neq \mathbf{0}$ esetben pedig válasszuk a legjobb javító irányt, azaz $\mathbf{d}_k = -\boldsymbol{\nabla} f(\mathbf{x}_k)$ és ekkor az eljárást a következő (4.) ponttal folytatjuk.

   (b) Ha van $\mathbf{M}$ mátrix és ennek sorvektorai lineárisan függetlenek, akkor legyen $\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}$ és legyen $\mathbf{d}_k = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}_k)$.

      i. Ha $\mathbf{d}_k \neq \mathbf{0}$, akkor az eljárást a következő ponttal folytatjuk.

      ii. Ha $\mathbf{d}_k = \mathbf{0}$, akkor kiszámítjuk a $\mathbf{w} = (\mathbf{u}, \mathbf{v}) = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\boldsymbol{\nabla} f(\mathbf{x}_k)$ Lagrange szorzók vektorát, ami $\mathbf{A}_1$ és $\mathbf{B}$ szerint partícionált.

         α Ha $\mathbf{u} \geqq \mathbf{0}$, akkor megállunk és az $\mathbf{x}_k$ megoldás KKT pont.

         β Ha $\mathbf{u} \ngeqq \mathbf{0}$, akkor kiválasztjuk az $\mathbf{u}$ vektor egyik negatív komponensét, legyen ez az u_r. Az $\mathbf{A}_1$ mátrixból elhagyjuk az r-edik sort és ezzel az új $\mathbf{A}_1$ mátrix-szal kiszámítjuk az új $\mathbf{M}$ mátrixot. Megismételjük az eljárás 3. pontját.

4. Megoldjuk a $\mathbf{d}_k$ javító lehetséges irányban az iránymenti optimalizálási feladatot, amely az alábbi:

$$\begin{aligned} \varphi(\lambda) &= f(\mathbf{x}_k + \lambda\mathbf{d}_k) \to \min! \\ 0 &\leqq \lambda \leqq \lambda_{\max} \end{aligned}$$

   ahol

$$\lambda_{\max} = \begin{cases} \min\left\{\frac{(\mathbf{b}_2 - \mathbf{A}_2\mathbf{x}_k)_i}{(\mathbf{A}_2\mathbf{d}_k)_i} : (\mathbf{A}_2\mathbf{d}_k)_i > 0\right\}, & \text{ha } \mathbf{A}_2\mathbf{d}_k \nleqq \mathbf{0} \\ \infty, & \text{ha } \mathbf{A}_2\mathbf{d}_k \leqq \mathbf{0} \end{cases}$$

   Legyen az optimális megoldás λ_k.

5. Meghatározzuk a következő közelítést: $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k\mathbf{d}_k$.

6. Folytatjuk az eljárást, $k := k + 1$.

10. Példa

Oldjuk meg a 6. példában szereplő optimalizálási feladatot Rosen féle gradiens projekciós módszerrel!

$$\begin{aligned} x_1^2 + x_2^2 &\to \min! \\ x_1 - x_2 &\leqq 2 \\ -x_1 - x_2 &\leqq -4 \end{aligned}$$

Megoldás

A célfüggvény gradiense $\nabla f(\mathbf{x}) = (2x_1, 2x_2)$. A példában $\mathbf{B}$ mátrix nem szerepel az $\mathbf{A}$ mátrix, ill. a $\mathbf{b}$ vektor pedig az alábbi

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}.$$

$k = 1$

Legyen az $\mathbf{x}_1 = (5,3)$ egy lehetséges megoldás. A célfüggvény gradiense az $\mathbf{x}_1 = (5,3)$ pontban $\nabla f(\mathbf{x}_1) = (10,6)$. Az első feltétel aktív, a második feltétel inaktív. Az $\mathbf{M}$ mátrix most egy sorból áll, mégpedig az $\mathbf{A}$ mátrix első sorvektora lesz:

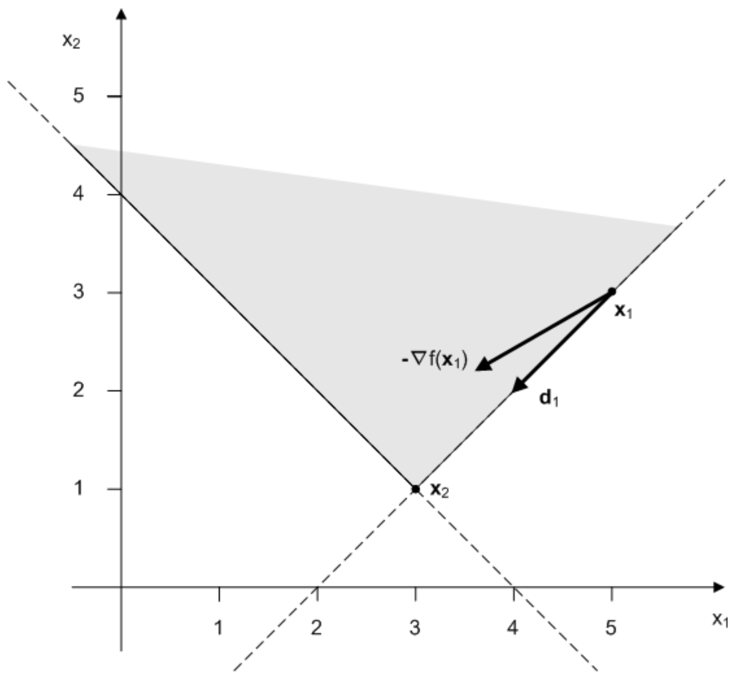
$$\mathbf{M} = \begin{bmatrix} 1 & -1 \end{bmatrix}.$$

A $\mathbf{P}$ projekciós mátrix meghatározása:

$$\begin{aligned} \mathbf{P} &= \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M} = \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -1 \end{bmatrix}^T \left(\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix}^T \right)^{-1} \begin{bmatrix} 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \end{aligned}$$

A projekciós mátrix ismeretében a javító lehetséges irány meghatározása: $\mathbf{d}_1 = -\mathbf{P}\nabla f(\mathbf{x}_1) = (-8,-8)$. Egyszerűbb számolás miatt használjuk a $\mathbf{d}_1 = (-1,-1)$ irányvektort.

Az alábbi ábrán nyomon követhetjük az eddigieket.



Megfigyelhetjük, hogy a $-\nabla f(\mathbf{x}_1)$ vektort a $\mathbf{P}$ projekciós mátrix az aktív feltételbe vetítette, így kaptuk a $\mathbf{d}_1$ irányt. A $\mathbf{d}_1 = (-1,-1)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_1 + \lambda\mathbf{d}_1 = (5,3) + \lambda(-1,-1) = (5-\lambda, 3-\lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\lambda_{\max} = \min\left\{ \frac{-4-(-5-3)}{2} \right\} = 2.$$

Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (5-\lambda)^2 + (3-\lambda)^2 \to \min! \\ 0 &\leqq \lambda \leqq 2 \end{aligned}$$

Ennek optimális megoldása $\lambda_1 = 2$, így a következő pont $\mathbf{x}_2 = \mathbf{x}_1 + \lambda_1 \mathbf{d}_1 = (3, 1)$.

$k = 2$

Az $\mathbf{x}_2 = (3, 1)$ lehetséges megoldásnál az első feltétel is és a második feltétel is aktív. A célfüggvény gradiense az $\mathbf{x}_2 = (3, 1)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_2) = (6, 2)$. Az $\mathbf{M}$ mátrix most maga az $\mathbf{A}$ mátrix

$$\mathbf{M} = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}.$$

A $\mathbf{P}$ projekciós mátrix meghatározása:

$$\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

A projekciós mátrix ismeretében a $\mathbf{d}$ vektor meghatározása: $\mathbf{d}_2 = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}_2) = (0, 0)$. A $\mathbf{d}_2$ zérusvektor, így kiszámítjuk a $\mathbf{w} = \mathbf{u}$ Lagrange szorzókat az alábbi képlettel

$$\mathbf{w} = \mathbf{u} = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\boldsymbol{\nabla} f(\mathbf{x}_2) = \begin{bmatrix} -2 \\ 4 \end{bmatrix}.$$

Mivel nem minden Lagrange szorzó nemnegatív, így az $\mathbf{x}_2 = (3, 1)$ pont nem KKT pont, tovább kell folytatni az eljárást. Az $u_1 < 0$, ezért az első sort kell elhagyni az aktív feltételek közül, azaz a módosított $\mathbf{M}$ mátrix (a példamegoldásban nem használjuk a tételben szereplő $\hat{\mathbf{M}}$ jelölést)) az alábbi

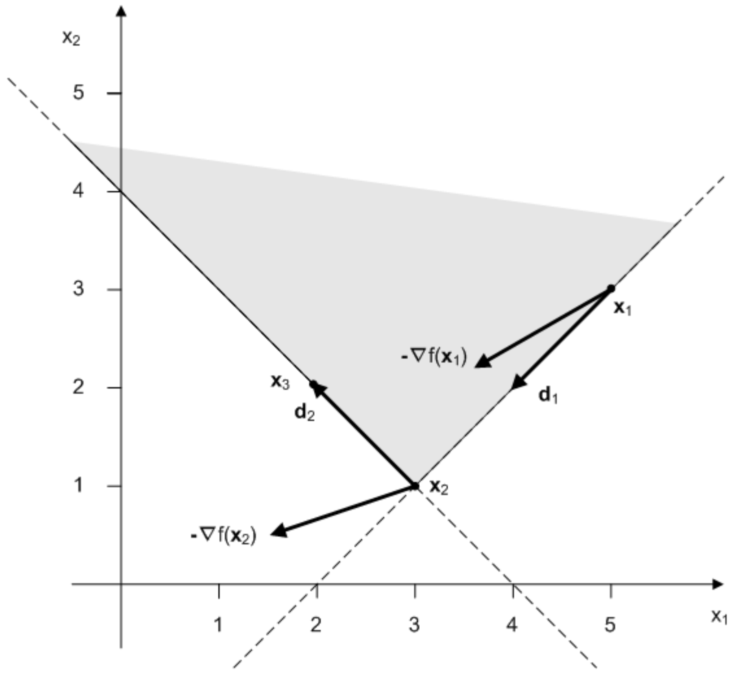
$$\mathbf{M} = \begin{bmatrix} -1 & -1 \end{bmatrix}.$$

Most ehhez az $\mathbf{M}$ mátrixhoz határozzuk meg a projekciós mátrixot, amely az alábbi

$$\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Az új projekciós mátrix ismeretében az új javító lehetséges irány meghatározása: $\mathbf{d}_2 = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}_2) = (-2, 2)$.

Az alábbi ábrán nyomon követhetjük az eddigieket.



Most is megfigyelhetjük, hogy a $-\boldsymbol{\nabla} f(\mathbf{x}_2)$ vektort a $\mathbf{P}$ projekciós mátrix az aktív feltételbe vetítette, így kaptuk a $\mathbf{d}_2$ irányt. Megjegyezzük, hogy ennél a lépésnél a legjobb irány a $-\boldsymbol{\nabla} f(\mathbf{x}_2)$ nem is lehetséges irány, de a vetítése már lehetséges irányt adott. Az ábrákból látható, hogy az első lépésnél a legjobb irány (a $-\boldsymbol{\nabla} f(\mathbf{x}_1)$ vektor) lehetséges irány, de az algoritmus nem figyeli ezt, az algoritmus mindig vetít, így mindig lehetséges irányt kapunk. Egyszerűbb számolás miatt használjuk a $\mathbf{d}_2 = (-1, 1)$ irányvektort. A $\mathbf{d}_2 = (-1, 1)$ irányban az új $\mathbf{x}$ vektor, $\mathbf{x} = \mathbf{x}_2 + \lambda \mathbf{d}_2 = (3, 1) + \lambda(-1, 1) = (3 - \lambda, 1 + \lambda)$. A $\lambda_{\max}$ értékének meghatározása

$$\lambda_{\max} = \infty.$$

Az egyváltozós iránymenti optimalizálási feladat

$$\begin{aligned} \varphi(\lambda) &= (3-\lambda)^2 + (1+\lambda)^2 \to \min! \\ \lambda &\geqq 0 \end{aligned}$$

Ennek optimális megoldása $\lambda_2 = 1$, így a következő pont $\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2 \mathbf{d}_2 = (2, 2)$.

$k = 3$

Az $\mathbf{x}_3 = (2, 2)$ lehetséges megoldásnál az első feltétel inaktív, a második feltétel aktív. A célfüggvény gradiense az $\mathbf{x}_3 = (2, 2)$ pontban $\boldsymbol{\nabla} f(\mathbf{x}_3) = (4, 4)$. Az $\mathbf{M}$ mátrix

$$\mathbf{M} = \begin{bmatrix} -1 & -1 \end{bmatrix}.$$

A $\mathbf{P}$ projekciós mátrix

$$\mathbf{P} = \mathbf{E} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

A projekciós mátrix ismeretében a $\mathbf{d}$ vektor meghatározása: $\mathbf{d}_3 = -\mathbf{P}\boldsymbol{\nabla} f(\mathbf{x}_3) = (0, 0)$. Mivel a $\mathbf{d}_3$ vektor zérus, így kiszámítjuk a Lagrange szorzókat

$$\mathbf{w} = \mathbf{u} = -(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M}\boldsymbol{\nabla} f(\mathbf{x}_3) = 4.$$

Mivel a Lagrange szorzó pozitív, így az $\mathbf{x}_3 = (2, 2)$ pont KKT pont és ezzel befejeztük az eljárást.

Megjegyzés:

Ezt a példát háromféle módszerrel is megoldottuk. Egyrészt a KKT pont tárgyalását tartalmazó tananyagban (meghatároztuk a feladat KKT pontját is), másrészt ebben a tananyagban Zoutendijk módszerrel és most Rosen-féle gradiens projekciós módszerrel. Javasoljuk az olvasónak a módszerek összehasonlítását. A Lagrange szorzókat is ellenőrizhetjük mind az $\mathbf{x}_2 = (3, 1)$, mind az $\mathbf{x}_3 = (1, 3)$ pont esetében.

3.3.3. Nemlineáris feltételek esete

Most olyan optimalizálási feladattal foglalkozunk, amelyben a feltételek nem szükségképpen lineárisak.

A feltételes optimalizálási feladat alakja a szokásos formában

$$\begin{aligned} f(\mathbf{x}) &\rightarrow \min! \\ g_i(\mathbf{x}) &\leqq 0, \quad i = 1, \ldots, m \\ h_i(\mathbf{x}) &= 0, \quad i = 1, \ldots, k \end{aligned}$$

A nemlineáris esetben az algoritmus hasonlóan működik, mint a fentebb leírt lineáris esetben. Legyen $\mathbf{x}_k$ egy lehetséges megoldás, azaz fennál, hogy $g_i(\mathbf{x}_k) \leqq 0, \quad i = 1, \ldots, m$ és $h_i(\mathbf{x}_k) = 0, \quad i = 1, \ldots, k$. Legyen az aktív egyenlőtlenséges feltételek indexhalmaza $I = \{i : g_i(\mathbf{x}) = 0\}$. Ekkor az $\mathbf{M}$ mátrix sorvektorai az aktív egyenlőtlenséges feltételek és az egyenlőséges feltételek gradiensei lesznek. Mint tudjuk, az $\mathbf{M}$ mátrix segítségével a megismert képlettel számított $\mathbf{P}$ projekciós mátrix a célfüggvény gradiensét a feltételi függvények gradiensvektorai által meghatározott nulltérbe viszi át. Ezért a $\mathbf{d} = -\mathbf{P}\nabla f(\mathbf{x}_k)$ irányvektor nem lesz lehetséges irány, mivel a lehetséges tartomány érintőjébe mutat. Amikor tehát az egyváltozós optimalizálást elvégezzük, akkor az új $\mathbf{x}_{k+1}$ nem lesz lehetséges megoldás, így egy korrekciót kell elvégezni, aminek eredményeképpen, már lehetséges megoldást kapunk. Ennek a korrekciónak az ismertetésére nem térünk ki.