Linjär programmering

Ett *LP-problem* (Linjärt Programmerings-problem) är ett optimeringsproblem där objektfunktionen och bivillkoren är linjära funktioner

Ett exempel på ett LP-problem är

$$\begin{array}{rcl} \max\limits_{x_1,x_2} & & x_1 + x_2 \\ & \text{s.t.} & \\ x_1 + 2x_2 & \leq & 8 \\ 3x_1 + 2x_2 & \leq & 12 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \end{array}$$

LP-problem

Vanligen skrivs objektfunktionen

$$f(x) = c^T x,$$

där c är en konstant vektor. Gradienten till $f(x)$ blir då

$$\nabla f(x) = c.$$

Definiera $F(\alpha)$ som funktionsvärdet utgående från en punkt $\bar{x}$ längs en sökriktning p och med steglängden α:

$$\begin{aligned} F(\alpha) &= f(\bar{x} + \alpha p) = c^T(\bar{x} + \alpha p) \\ &= c^T\bar{x} + \alpha c^T p = f(\bar{x}) + \alpha c^T p. \end{aligned}$$

Nedförsriktningar

Förändringen i funktionsvärdet beror alltså på $c^T p$. $c^T p = 0$ innebär att

$$F(\alpha) = f(\bar{x}) + \alpha \underbrace{c^T p}_{=0} = f(\bar{x})$$

förblir konstant. $c^T p < 0$ innebär att

$$F(\alpha) = f(\bar{x}) + \alpha \underbrace{c^T p}_{<0} < f(\bar{x})$$

för $\alpha > 0$, dvs funktionen minskar. Riktningen p kallas då för *nedförsriktning*.

Bivillkor

Vi har två olika sorters bivillkor: likhets- och olikhetsbivillkor. Vi kommer att koncentrera oss på olikhetsbivillkor.

Linjära olikhetsbivillkor kan skrivas som

$$a^T x \geq \beta$$

för någon vektor a. En punkt $\bar{x}$ sägs vara *tillåten* med avseende på villkoret $a^T x \geq \beta$ om villkoret är satisfierat, dvs $a^T \bar{x} \geq \beta$. Om villkoret är satisfierat med likhet ($a^T \bar{x} = \beta$) sägs villkoret vara *aktivt*, annars *inaktivt*.

Två olikheter sägs vara *ekvivalenta* om de har identiska mängder av tillåtna punkter. Speciellt är $a^T x \leq \beta$ ekvivalent med $-a^T x \geq -\beta$, vilket innebär att samtliga $\leq$-bivillkor kan skrivas om till $\geq$-bivillkor.

Vi behöver därför endast resonera runt $\geq$-bivillkor.

Mängder av bivillkor

En mängd bivillkor $\{a_i^T x \geq \beta_i, i = 1, \ldots, m\}$ representeras vanligen av en $m \times n$-matris A och m-vektor b:

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}, b = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$$

och skrivs

$$Ax \geq b,$$

där $\geq$ appliceras komponentvis.

En punkt $\bar{x}$ sägs vara *tillåten* med avseende på mängden olikheter $Ax \geq b$ om $A\bar{x} \geq b$.

Systemet $Ax \geq b$ sägs vara *konsistent* om det har någon tillåten punkt.

Geometrisk tolkning av bivillkor

Varje bivillkor delar upp $\Re^n$ i två halvor av tillåtna punkter $\{x : a^T x \geq b\}$ och otillåtna punkter $\{x : a^T x < b\}$.

Definiera mängden F av tillåtna punkter som

$$\mathcal{F} \equiv \{x \in \Re^n : Ax \geq b\}.$$

Om $Ax \geq b$ är konsistent kan $\mathcal{F}$ vara finit (sluten) eller infinit (öppen). I bägge fallen är $\mathcal{F}$ konvex.

Aktiva och inaktiva mängder av bivillkor

För ett villkor $a_i^T x \geq \beta_i$ definieras *residualen* i punkten $\bar{x}$ som $r_i(\bar{x}) = a_i^T \bar{x} - \beta_i$. Residualen är positiv då villkoret är inaktivt, noll då villkoret är aktivt och negativt då villkoret ej är uppfyllt.

Residualvektorn r blir för mängden av villkor $Ax \geq b$

$$r(\bar{x}) = A\bar{x} - b$$

Givet olikheten $Ax \geq b$ definierar vi indexmängden $\mathcal{A}(\bar{x})$ som mängden av index $\{i : r_i(\bar{x}) = 0\}$, dvs de villkor som uppfylls med likhet i punkten $\bar{x}$. Matrisen bestående av raderna i $\mathcal{A}(\bar{x})$ betecknas $A_{\mathcal{A}}(\bar{x})$.

Förflyttning längs en riktning

Studera en (icke nödvändigtvis tillåten) punkt $\bar{x}$, en riktning p och ett villkor $a_i^T x \geq \beta_i$. För ett steg α längs p gäller att

$$a_i^T(\bar{x} + \alpha p) = a_i^T \bar{x} + \alpha a_i^T p.$$

Motsvarande residual blir

$$r_i(\bar{x} + \alpha p) = a_i^T \bar{x} + \alpha a_i^T p - \beta_i = r_i(\bar{x}) + \alpha a_i^T p.$$

Om $a_i^T p = 0$ förändras inte värdet på residualen, men för $a_i^T p \neq 0$ går det att räkna ut hur långt steg σ_i vi ska ta för att satisfiera och aktivera bivillkoret:

$$\begin{aligned} r_i(\bar{x} + \sigma_i p) &= 0 \\ r_i(\bar{x}) + \sigma_i a_i^T p &= 0 \\ &\Downarrow \\ \sigma_i &= \frac{r_i(\bar{x})}{-a_i^T p} \end{aligned}$$

som är definierat då $a_i^T p \neq 0$. Då $a_i^T p = 0$ definieras σ_i som ∞ om $r_i(\bar{x}) \geq 0$ och $-\infty$ annars.

Tillåtna riktningar

Vektorn $p \neq 0$ är en *tillåten riktning* i den tillåtna punkten $\bar{x}$ med avseende på villkoret $a_i^T x \geq \beta_i$ om det existerar en *positiv* skalär τ_i sådan att

$$r_i(\bar{x} + \alpha p) \geq 0 \;,\; 0 \leq \alpha \leq \tau_i.$$

För ett inaktivt bivillkor ($r_i(\bar{x}) > 0$) ser vi att alla riktningar p är tillåtna, medan för ett aktivt bivillkor ($r_i(\bar{x}) = 0$) gäller att

$$\begin{aligned} r_i(\bar{x} + \alpha p) &\geq 0 \\ &\Downarrow \\ a_i^T(\bar{x} + \alpha p) - \beta_i &= r_i(\bar{x}) + \alpha a_i^T p \\ &= \alpha a_i^T p \geq 0,\; \alpha > 0 \\ &\Updownarrow \\ a_i^T p &\geq 0 \end{aligned}$$

Vi ser att om $a_i^T p = 0$ så förblir villkoret aktivt, medan $a_i^T p > 0$ leder till att villkoret blir inaktivt och $a_i^T p < 0$ leder till att villkoret inte längre uppfylls.

Tillåtna riktningar, forts.

För en mängd av bivillkor gäller att p är en tillåten riktning från den tillåtna punkten $\bar{x}$ om $p \neq 0$ och det existerar en positiv skalär τ sådan att

$$\begin{aligned} r(\bar{x} + \alpha p) &\geq 0 \;,\; 0 \leq \alpha \leq \tau \\ &\Updownarrow \\ A(\bar{x} + \alpha p) &\geq b \;,\; 0 \leq \alpha \leq \tau \end{aligned}$$

Då de enda begränsningarna på p kommer från de aktiva bivillkoren, gäller att p är tillåten omm

$$\begin{aligned} a_i^T p &\geq 0 \;\; \forall i \in \mathcal{A}(\bar{x}) \\ &\Updownarrow \\ A_{\mathcal{A}} p &\geq 0, \end{aligned}$$

där $A_{\mathcal{A}}$ betecknar den aktiva mängdmatrisen i $\bar{x}$.

Tillåtna nedförsriktningar

Studera en konsistent villkorsmängd $Ax \geq b$ och en tillåten punkt x^*. Om det existerar en tillåten nedförsriktning i x^* så kan x^* inte vara en minpunkt.

En riktning $p \neq 0$ är tillåten om $A_{\mathcal{A}} p \geq 0$ och nedförsriktning om $c^T p < 0$.

Omm ingen sådan riktning existerar är x^* en minpunkt. Detta inträffar om $\lambda \geq 0$, där λ är lösningen till

$$A_{\mathcal{A}}^T \lambda = c.$$

Val av sökriktning

Om vi löser

$$A_{\mathcal{A}}^T \lambda = c$$

och något λ_i är negativt, så kan en sökriktning p bestämmas som lösningen till

$$A_{\mathcal{A}} p = e_i,$$

där e_i är en vektor med i:te elementet 1 och resten 0. Det innebär att $r_i(\bar{x} + \alpha p) > 0$, dvs villkor i blir inaktiverat.

Väljes p på detta sätt blir p en nedförsriktning, ty $c^T p = (A_{\mathcal{A}}^T \lambda)^T p = \lambda^T A_{\mathcal{A}} p = \lambda^T e_i = \lambda_i < 0$.

Hörn

Ett *hörn* är en extrempunkt i den tillåtna mängden som inte kan ligga på någon linje mellan två andra tillåtna punkter. Ett hörn är begränsat i alla n dimensioner, dvs den aktiva mängdmatrisen $A_{\mathcal{A}}$ har minst n linjärt oberoende rader. Om $A_{\mathcal{A}}$ har fler än n rader kallas hörnet för *degenererat*, annars kallas det för *icke-degenererat*.

Det går att visa att minimum för ett LP-problem antas i ett hörn.

Steglängd till nästa hörn

Antag att $\bar{x}$ är ett hörn i den tillåtna mängden och vi tagit ut en sökriktning p från $\bar{x}$. Låt mängden $\mathcal{D}$ beteckna mängden av alla inaktiva villkor som minskar längs p, dvs

$$j \in \mathcal{D} \text{ om } a_j^T p < 0.$$

Om vi bestämmer σ_j som det längsta steg som kan tas längs med p utan att villkor j aktiveras dvs,

$$\sigma_j = \begin{cases} \frac{-r_j(\bar{x})}{a_j^T p} & \text{om } j \in \mathcal{D} \\ +\infty & \text{annars} \end{cases},$$

så blir

$$\alpha = \min_j \sigma_j$$

det längsta steg vi kan ta innan något bivillkor aktiveras. Den nya punkten blir då $\bar{x} + \alpha p$.

Simplex-metoden

Låt $x^{[n]}$ beteckna den aktuella punkten efter n iterationer. Givet en tillåten punkt $x^{[n]}$ så blir algoritmen för simplex-metoden:

- Bestäm en tillåten nedförsriktning p i $x^{[n]}$:
  - Identifiera den aktiva mängden $A_{\mathcal{A}}$ i $\bar{x}$.
  - Lös $A_{\mathcal{A}}^T \lambda = c$.
  - Välj i sådan att $\lambda_i < 0$.
  - Bestäm sökriktningen p som lösningen till $A_{\mathcal{A}} p = e_i$.
- Bestäm steglängden $\alpha = \min_j \sigma_j$, där

$$\sigma_j = \begin{cases} \frac{-r_j(\bar{x})}{a_j^T p} & \text{om } j \in \mathcal{D} \\ +\infty & \text{annars} \end{cases}.$$

- $x^{[n+1]} = x^{[n]} + \alpha p$.