Linjära bivillkor

Ett fullständigt optimeringsproblem med linjära bivillkor kan skrivas

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ \text{s.t.} & a_i^T x = b_i, \quad i \in \mathcal{E} \\ & a_i^T x \geq b_i, \quad i \in \mathcal{I} \end{array}$$

Vid beräkning av en tillåten stegriktning p i en punkt $\bar{x}$ styrs denna av de aktiva bivillkoren i $\bar{x}$. Vi ska nu studera likhetsbivillkoren lite närmare.

Nollrum och värderum

Låt A vara en $m \times n$-matris med $m \leq n$. Vi definierar *nollrummet* till A som

$$\mathcal{N}(A) = \{p \in \Re^n : Ap = 0\}.$$

Dimensionen hos $\mathcal{N}(A)$ är $n-r$, där r är rangen hos A^T.

Om A innehåller likhetsbivillkoren för ett problem, så innehåller $\mathcal{N}(A)$ alla tillåtna sökriktningar.

Vi definierar också *värderummet* till en matris A^T som

$$\mathcal{R}(A^T) = \{q \in \Re^n : q = A^T\lambda \text{ för något } \lambda \in \Re^m\}.$$

Dimensionen hos $\mathcal{R}(A^T)$ är r.

Nollrum och värderum, forts.

Rummen $\mathcal{N}(A)$ och $\mathcal{R}(A^T)$ är ortogonala mot varandra, då för varje vektor $q \in \mathcal{R}(A^T)$ och $p \in \mathcal{N}(A)$ gäller att

$$q^T p = (A^T\lambda)^T p = \lambda^T A p = 0$$

för något $\lambda \in \Re^m$.

Då deras dimensioner dessutom summeras till n kan varje vektor $x \in \Re^n$ skrivas som en unik summa

$$x = p + q,$$

där $p \in \mathcal{N}(A)$ och $q \in \mathcal{R}(A^T)$.

Nollrumsmatris

Definiera matrisen Z som *nollrumsmatris* om $\mathcal{R}(Z) = \mathcal{N}(A)$. Om A har full radrang m så är varje $n \times r$-matris Z med rang $n-m$ som uppfyller $AZ = 0$ en nollrumsmatris. Om $r = n - m$ har Z linjärt oberoende kolumner och sägs vara en *bas* för $\mathcal{N}(A)$. Då kan nollrummet representeras som

$$\mathcal{N}(A) = \{p : p = Zv \text{ för något } v \in \Re^r\}.$$

Nollrumsmatris, forts.

Med hjälp av detta kan vi lätt beskriva alla tillåtna punkter: Om $\bar{x}$ är en punkt som uppfyller $Ax = b$ så kan alla andra tillåtna punkter beskrivas som

$$x = \bar{x} + Zv$$

för någon vektor v, dvs bivillkorsproblemet

$$\begin{array}{ll} \min\limits_{x} & f(x) \\ \text{s.t.} & Ax = b \end{array}$$

går att formulera om till

$$\min_{v} \quad f(\bar{x} + Zv),$$

vilket är ett problem *utan* bivillkor.

Konstruktion av nollrumsmatriser

Studera att bivillkor $Ap = 0$, där A är $m \times n$.

Antag att de m första kolumnerna i A är linjärt oberoende. (Annars kan alltid A och p permuteras så att de blir det.) Då får vi

$$Ap = [B\ N] \begin{bmatrix} p_B \\ p_N \end{bmatrix} = Bp_B + Np_N = 0.$$

Om vi multiplicerar med B^{-1} från vänster får vi $p_B = -B^{-1}Np_N$, dvs lösningsmängden till $Ap = 0$ är

$$p = \begin{bmatrix} p_B \\ p_N \end{bmatrix} = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} p_N$$

och $n \times (n-m)$-matrisen

$$Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix}$$

utgör en bas för $\mathcal{N}(A)$.

Konstruktion av nollrumsmatriser, forts.

Studera nu problemet $Ax = b$. En tillåten punkt är

$$\bar{x} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix},$$

ty

$$A\bar{x} = [B\ N] \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} = BB^{-1}b + N \cdot 0 = b.$$

Det innebär att alla punkter x som uppfyller $Ax = b$ kan skrivas som

$$x = \bar{x} + p = \bar{x} + Zp_N = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} + \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} p_N.$$

Detta innebär att vi reducerat problemet från n parametrar till $n-m$ parametrar och kallas därför för variabelreduktionsmetoden.

Ortogonala projektionsmatriser

Låt x vara en n-vektor och A en $m \times n$-matris med full radrang m. Då kan x skrivas som

$$x = p + q \Leftrightarrow p = x - q,$$

där $Ap = 0$ och $q = A^T\lambda$ för någon m-vektor λ. Om vi multiplicerar med A från vänster får vi

$$Ax = Ap + Aq = AA^T\lambda \Rightarrow \lambda = (AA^T)^{-1}Ax.$$

Insättning ger nollrumskomponenten

$$\begin{aligned} p &= x - A^T\lambda = x - A^T(AA^T)^{-1}Ax \\ &= (I - A^T(AA^T)^{-1}A)x, \end{aligned}$$

där $n \times n$-matrisen

$$P = I - A^T(AA^T)^{-1}A$$

kallas för den *ortogonala projektionsmatrisen* på $\mathcal{N}(A)$.

Nollrumskomponenten av en vektor x fås genom att multiplicera med P. Px kallas också för den ortogonala projektionen av x på $\mathcal{N}(A)$.

Ortogonala projektionsmatriser, forts.

Den ortogonala projectionsmatrisen P uppfyller:

- P är en nollrumsmatris till A; $AP = 0$.
- $P^2 = P$, dvs en vektor $v \in \mathcal{N}(A)$ påverkas inte av projektionen,
- $P = P^T$, dvs P är symmetrisk.

Notera att projektionsmatrisen P sällan bildas, utan projektionen Px beräknas som $Px = x - A^T(AA^T)^{-1}Ax$. Speciellt om A består av en enda rad a, så blir $Px = x - \frac{1}{a^T a} a(a^T x)$ mycket billigt att beräkna.

QR-faktoriseringen

Låt A vara en $m \times n$-matris med full radrang m. Om vi utför en ortogonal faktorisering

$$A^T = QR = [Q_1 \; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix},$$

där Q_1 är $n \times m$, Q_2 är $n \times (n-m)$ och R_1 är $m \times m$.

Då Q är en ortogonal matris blir

$$\begin{aligned} Q^T A^T &= Q^T Q R = R \\ &\Updownarrow \\ AQ &= R^T = \begin{bmatrix} R_1^T & 0 \end{bmatrix} \\ &\Updownarrow \\ AQ_1 &= R_1^T \text{ och } AQ_2 = 0, \end{aligned}$$

alltså är Q_2 en bas för $\mathcal{N}(A)$. Q_2 är också en *ortogonal* bas, ty $Q_2^T Q_2 = I$.

För icke stora, eller icke glesa problem är QR-faktoriseringen att föredra då den är numeriska stabil.