KTH Matematik

5B1817 Tillämpad ickelinjär optimering

Föreläsning 6

Kvadratisk programmering med likhetsbivillkor

Kvadratisk programmering med likhetsbivillkor

Titta på modellproblem med kvadratisk målfunktion,

$$
(EQP) \quad \begin{array}{ll} \text{min} & f(x) = \frac{1}{2}x^T H x + c^T x \\ \text{då} & Ax = b, \\ & x \in I\!R^n. \end{array}
$$

Vi antar att $A \in I\!R^{m \times n}$ med rang m.

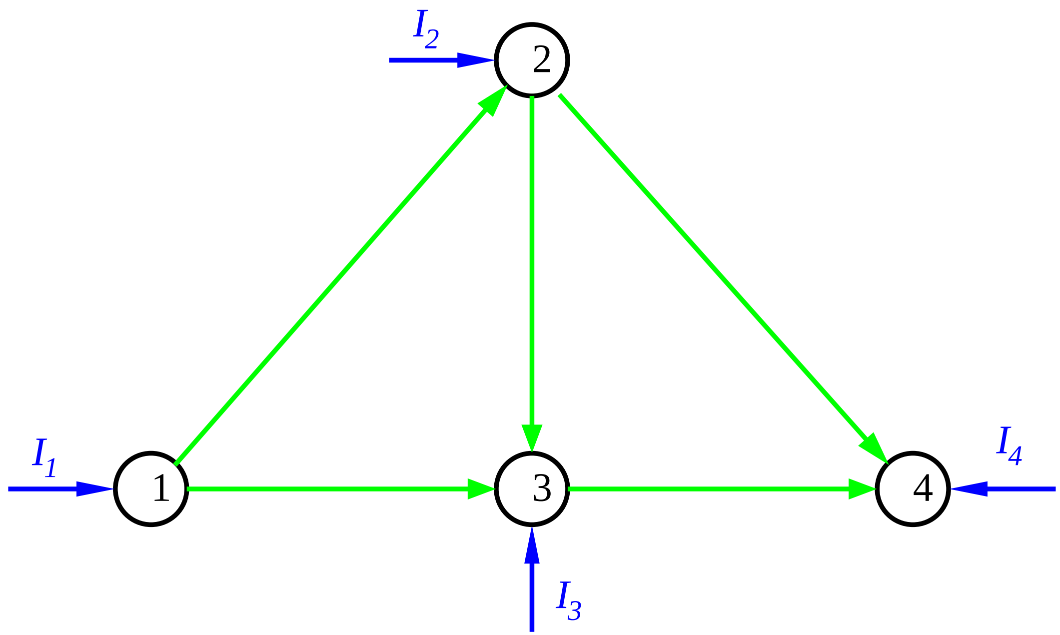
Första ordningens nödvändiga optimalitetsvillkor blir

$$
\begin{aligned} Hx + c &= A^T \lambda, \\ Ax &= b. \end{aligned}
$$

Detta är ett linjärt ekvationssystem.

Exempelproblem, elektriskt nätverk

Antag att vi har ett elektiskt nätverk med m noder och ledningar $(i, j) \in E$. Ström med styrka I_i skickas in i nod i, $i = 1, \dots, m$ (där $\sum_{i=1}^{m} I_i = 0$). Ledning (i, j) har resistans r_{ij}.



Vilken väg tar strömmen genom nätverket?

Exempelproblem, elektriskt nätverk, forts.

Låt x_{ij} vara strömstyrkan i ledning (i,j) och låt u_i vara potentialen i nod i, $i = 1,\ldots,m$, där vi kan välja $u_m = 0$. Då får vi

$$\begin{aligned} u_i - u_j &= r_{ij} x_{ij}, \quad (i,j) \in E, \\ \sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ji} &= I_i, \quad i = 1,\ldots,m-1. \end{aligned}$$

Ekvivalent med optimallösning och multiplikatorer till

$$\begin{array}{ll} \min & \frac{1}{2} \sum_{(i,j)\in E} r_{ij} x_{ij}^2 \\ \text{då} & \sum_{j:(i,j)\in E} x_{ij} - \sum_{j:(j,i)\in E} x_{ji} = I_i, \quad i = 1,\ldots,m-1, \\ & x_{ij} \in I\!R, \ (i,j) \in E. \end{array}$$

Totala effekten minimeras.

Optimalitetsvillkor för kvadratisk programmering med likhetsbivillkor

Första ordningens nödvändiga optimalitetsvillkor kan skrivas

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ -\lambda \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}.$$

Låt Z vara en matris vars kolumner bildar bas för null(A).

Påstående. *En punkt $x^* \in \mathbb{R}^n$ är en global minpunkt till (EQP) om och endast om det finns $\lambda^* \in \mathbb{R}^m$ så att*

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ -\lambda^* \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix} \quad \textit{och} \quad Z^T H Z \succeq 0.$$

Kvadratisk programmering med likhetsbivillkor

Alternativt, låt x vara given punkt och p steget till optimum,

$$(EQP') \quad \begin{array}{ll} \min & f(x+p) = \frac{1}{2}(x+p)^T H(x+p) + c^T(x+p) \\ \text{då} & Ap = b - Ax, \\ & p \in \mathbb{R}^n. \end{array}$$

Påstående. *En punkt $x + p^* \in \mathbb{R}^n$ är en global minpunkt till (EQP) om och endast om det finns $\lambda^* \in \mathbb{R}^m$ så att*

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p^* \\ -\lambda^* \end{pmatrix} = - \begin{pmatrix} Hx + c \\ Ax - b \end{pmatrix} \quad \textit{och} \quad Z^T H Z \succeq 0.$$

OBS! Samma λ^* som tidigare.

KKT-matrisen

Matrisen $K = \begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix}$ kallas *KKT-matrisen*.

Påstående. *Om* $A \neq 0$ *är* $K \not\succeq 0$.

Detta medför att K är en indefinit matris.

Påstående. *Om* $Z^T H Z \succ 0$ *och* $\text{rank}(A) = m$ *är* K *ickesingulär.*

Om $Z^T H Z \succ 0$ och $\text{rank}(A) = m$ blir x^* och λ^* alltså unika.

Vi antar för resten av dagen att $Z^T H Z \succ 0$ och $\text{rank}(A) = m$.

Hur räknar vi ut x^* och λ^*?

Vi föredrar (EQP') framför (EQP).

Direkt lösning med K

Vi kan räkna ut p^* och λ^* direkt med hjälp av K, dvs

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p^* \\ -\lambda^* \end{pmatrix} = - \begin{pmatrix} Hx + c \\ Ax - b \end{pmatrix}.$$

Vi behöver lösa ekvationssystem där matrisen är symmetrisk och indefinit.

Lösning med partitionerad K

Låt $p^* = Zp_Z^* + Yp_Y^*$, där $Y \in I\!R^{n\times m}$ väljs så att $(Z\ Y)$ är ickesingulär.

Exempelvis $Y = A^T$ eller $Y = \begin{pmatrix} I \\ 0 \end{pmatrix}$ där $A = \begin{pmatrix} B & N \end{pmatrix}$ med

$B \in I\!R^{m\times m}$, B ickesingulär.

Vi kan då skriva
$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} Z & Y & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} p_Z^* \\ p_Y^* \\ -\lambda^* \end{pmatrix} = - \begin{pmatrix} Hx + c \\ Ax - b \end{pmatrix}.$$

Multiplikation från vänster med $\begin{pmatrix} Z^T & 0 \\ Y^T & 0 \\ 0 & I \end{pmatrix}$ ger symmetrisk matris.

Lösning med partitionerad K, forts.

Symmetrisering genom multiplikation från vänster ger

$$\begin{pmatrix} Z^T & 0 \\ Y^T & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} Z & Y & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} p_Z^* \\ p_Y^* \\ -\lambda^* \end{pmatrix} =$$

$$- \begin{pmatrix} Z^T & 0 \\ Y^T & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} Hx + c \\ Ax - b \end{pmatrix}.$$

Observera att ekvationssystemen är ekvivalenta, då vi multiplicerar med ickesingulär matris.

Lösning med partitionerad K, forts.

Förenkling ger

$$\begin{pmatrix} Z^THZ & Z^THY & 0 \\ Y^THZ & Y^THY & Y^TA^T \\ 0 & AY & 0 \end{pmatrix} \begin{pmatrix} p_Z^* \\ p_Y^* \\ -\lambda^* \end{pmatrix} = - \begin{pmatrix} Z^T(Hx+c) \\ Y^T(Hx+c) \\ Ax-b \end{pmatrix}.$$

Ekvationssystemet kan nu lösas i tre steg:

(i) $AYp_Y^* = b - Ax,$ tillåten punkt

(ii) $Z^THZp_Z^* = -Z^T(H(x+Yp_Y^*)+c),$ optimal punkt

(iii) $Y^TA^T\lambda^* = Y^T(H(x+Yp_Y^*+Zp_Z^*)+c).$ lagrangemultiplikator

OBS! p_Y beror inte på H och c. Endast nollrumssteget p_Z gör det.

Nollrumsmetod

Antag speciellt att x är tillåten, dvs $Ax = b$. Då blir (i) trivial.

(i) $AYp_Y^* = 0$, dvs $p_Y^* = 0$ tillåten punkt

(ii) $Z^THZp_Z^* = -Z^T(Hx + c)$, optimal punkt

(iii) $Y^TA^T\lambda^* = Y^T(H(x + Zp_Z^*) + c)$. lagrangemultiplikator

Endast steget i nollrummet, p_Z, behöver bestämmas.

Den partitionerade metoden kallas därför ofta *nollrumsmetod*.

Om faktoriseringar av AY och Z^THZ är kända, kan (i), (ii) och (iii) lösas.

Val av metod

Alla metoder som löser (EQP) löser ekvationssystemet

$$\begin{pmatrix} Hx + c - A^T\lambda \\ Ax - b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Vi har tittat på två olika sätt att göra detta på.

Vilket sätt som är bäst beror på många faktorer.

En direkt metod kan utnyttja struktur i K, men måste hantera en indefinit matris.

En nollrumsmetod jobbar med matriser av mindre dimension. Det är dock svårare att ta hänsyn till strukturen i K.

Alternativt synsätt

Vi kan alternativt direkt betrakta det linjära ekvationssystemet

$$\begin{pmatrix} Hx + c - A^T\lambda \\ Ax - b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

För en given initialgissning x söker vi $x + p^*$ och $0 + \lambda^*$ så att

$$\begin{pmatrix} H(x + p^*) + c - A^T\lambda^* \\ A(x + p^*) - b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

dvs

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} p^* \\ -\lambda^* \end{pmatrix} = - \begin{pmatrix} Hx + c \\ Ax - b \end{pmatrix}.$$

Exempel, minstakvadratproblem

Betrakta ett minstakvadratproblem

$$\begin{array}{ll} \min & \|A^T y - c\|_2^2 \\ \text{då} & y \in \mathbb{R}^m. \end{array}$$

Normalekvationerna kan skrivas $AA^T y^* = Ac$.

Ekvivalent omskrivning $\begin{pmatrix} I & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -r^* \\ -y^* \end{pmatrix} = - \begin{pmatrix} c \\ 0 \end{pmatrix}$.

Ekvivalent problem

$$\begin{array}{ll} \min & \|r - c\|_2^2 \\ \text{då} & Ar = 0, \\ & r \in \mathbb{R}^n. \end{array}$$

Observation relaterat till problem med olikhetsbivillkor

Antag att $x^* = x + p^*$ med tillhörande λ^* är optimallösning till

$$\begin{array}{ll} \min & \frac{1}{2}x^T Hx + c^T x \\ \text{då} & Ax = b, \\ & x \in I\!R^n, \end{array}$$

där $H \succ 0$. Om $\lambda^* \geq 0$ är x^* då också optimallösning till

$$\begin{array}{ll} \min & \frac{1}{2}x^T Hx + c^T x \\ \text{då} & Ax \geq b, \\ & x \in I\!R^n. \end{array}$$

Denna observation är basen för en *active set* metod för att lösa kvadratiska programmeringsproblem med olikhetsbivillkor.