Kapitel 1

Quadratische Optimierung

1. Aufgabenstellung

Unter der Aufgabe der quadratischen Optimierung verstehen wir folgendes Problem:
Es sei die quadratische konvexe[1] Funktion

$$(1.\ 1.) \qquad Q(x) = \sum_{j=1}^{n}\sum_{k=1}^{n} x_j C_{jk} x_k + \sum_{j=1}^{n} p_j x_j$$

zu minimieren[2] unter den linearen Restriktionen

$$(1.\ 2.) \qquad f_i(x) \equiv \sum_{j=1}^{n} a_{ij} x_j - b_i \le 0, \qquad i = 1,2,...,m$$
$$x_j \ge 0, \qquad j = 1,2,...,n\,.$$

Anmerkung:

1. In Matrizenschreibweise lassen sich (1. 1.) und (1. 2.) verkürzt wie folgt schreiben:

$$(PI) \qquad Min\left\{Q(x) = x^T Cx + p^T x \mid Ax \le b,\ x \ge 0\right\}.$$

Hierbei sind:

$C = (c_{jk})$: eine positiv semidefinite[3] symmetrische $(n \cdot n)$-Matrix,
$A = (a_{ij})$: eine $(m \cdot n)$-Matrix,
$p = \left(p_1, p_2, ..., p_n\right)^T$,
$x = \left(x_1, x_2, ..., x_n\right)^T$,
$b = \left(b_1, b_2, ..., b_m\right)^T$.

2. Durch formale Umformungen kann man (*PI*) in die folgenden zwei Schreibweisen überführen:

$$(PII) \qquad Min\left\{Q(x) = x^T Cx + p^T x \mid Ax = b,\ x \ge 0\right\},$$

---

[1] Für das allgemeine quadratische Programm mit beliebiger (weder konvexe noch konkave) Zielfunktion existieren bisher weder eine umfassende noch zufriedenstellenden Algorithmen [45].

[2] In der ganzen Arbeit beschränken wir uns auf den Fall der Minimierung. Falls die Zielfunktion zu maximieren ist, minimiert man die Funktion $\{-Q(x)\}$. In diesem Falle wird die Zielfunktion offensichtlich konkav sein.

[3] $x^T Cx$ ist konvex dann und nur dann, wenn C positiv semidefinit ist.

(*PIII*) $\qquad Min\left\{Q(x) = x^T Cx + p^T x \mid Ax \le b\right\}.$

Die Probleme (*PI*), (*PII*) und (*PIII*) werden häufig als „*Normalformen*" der quadratischen Optimierung bezeichnet. Es zeigt sich jedoch, dass es bei verschiedenen Algorithmen zweckmäßiger ist diese oder jene Form von Restriktionen zu benutzen.

Satz 1 (*Über die Menge der optimalen Lösungen*):
Es sei x^* eine optimale Lösung des Problems der quadratischen Optimierung. Die zulässige Lösung $\bar{x}$ ist optimal dann und nur dann, wenn für den Vektor $\bar{x} - x^* = t$ gilt:

(1. 3.) $\qquad Ct = 0,$

(1. 4.) $\qquad p^T t = 0.$

Folgerung:
Hat das Gleichungssystem $Ct = 0,\ p^T t = 0$ keine nichttriviale Lösungen (z.B. wenn C streng positiv definit ist), so hat das Problem der quadratischen Optimierung höchstens eine optimale Lösung.

2. Optimalitätsbedingungen

Führt man für die Probleme (*PI*), (*PII*) und (*PIII*) die Lagrange-Funktion

(1. 5.) $\qquad L(x,u) = x^T Cx + p^T x + u^T (Ax - b)$

und die Hilfsvariablen

(1. 6.) $\qquad v = \dfrac{\partial L(x,u)}{\partial x} = p + 2Cx + A^T u$

und

(1. 7.) $\qquad y = -\dfrac{\partial L(x,u)}{\partial u} = -Ax + b$

ein, so lauten die Kuh-Tucker-Bedingungen[4]:

P(I):

(1. 8.)
$$\begin{cases} Ax + y = b & (A) \\ 2Cx + A^T u - v = -p & (B) \\ x \ge 0,\ u \ge 0,\ v \ge 0,\ y \ge 0 & (C) \\ x^T v + u^T y = 0 & (D) \end{cases}$$

[4] Im Kapitel III werden diese Bedingungen für einen allgemeineren Fall bewiesen.

P(II):

$$(1.\ 9.)\qquad \begin{cases} Ax = b & (A) \\ 2Cx + A^T u - v = -p & (B) \\ x \geq 0,\ v \geq 0 & (C) \\ x^T v = 0 & (D) \end{cases}$$

P(III):

$$(1.\ 10.)\qquad \begin{cases} Ax + y = b & (A) \\ 2Cx + A^T u = -p & (B) \\ u \geq 0,\ y \geq 0 & (C) \\ u^T y = 0 & (D) \end{cases}$$

<u>Satz 2</u> (*Kuh-Tucker*):
Ein Punkt x^* dann und nur dann eine optimale Lösung des Problems (*PI*) bzw. (*PII*) bzw. (*PIII*) dar, wenn Vektoren $u^*,\ v^*, y^*$ existieren, für welche (1. 8.) bzw. (1. 9.) bzw. (1. 10) erfüllt sind.

Anmerkung:

1. Die Bedingungen *A* bis *C* (ohne *D*) sind ein notwendiges und hinreichendes Kriterium für die Existenz einer optimalen Lösung des Problems (*PI*) bzw. (*PII*) bzw. (*PIII*).
2. In den Kuhn-Tucker-Bedingungen sind die Beziehungen *A* bis *C* linear, und die Beziehung *D* ist nichtlinear. Aus der Bedingung *D* folgt, dass von je zwei vorzeichenbeschränkten Variablen $x_j,\ v_j\ , j = 1,2,...,n,$ bzw. $u_i,\ y_j\ , i = 1,2,...,m,$ mindestens eine Komponente gleich Nuss sein muss, und somit können höchstens $m + n$ Variablen von Null verschieden sein. Die Bedingungen *A* und *B* bilden also ein lineares System mit $m + n$ Gleichungen, d.h., es sind höchstens so viel Variablen von Null verschieden, wie die Anzahl der Restriktionen im System *A* bis *C*. Die Tatsache, dass, wenn ein Gleichungssystem nichtnegative Lösungen hat, auch nichtnegative Basislösungen haben muss, ermöglicht es, sich bei der Untersuchung der Kuhn-Tucker-Bedingungen (*A-D*) nur auf die nichtnegativen Basislösungen des Systems *A* bis *B* zu beschränken (*Simplexprinzip*).

3. Das Dualproblem

Die Dualprobleme zu den Primalproblemen (*PI*), (*PII*) und (*PIII*) lassen sich in folgender Form darstellen:

$$(D1)\quad Max\left\{\Theta(x,u) = -x^T Cx - b^T u \mid 2Cx + A^T u \geq -p,\ u \geq 0\right\},$$

$$(DI1)\quad Max\left\{\Theta(x,u) = -x^T Cx - b^T u \mid 2Cx + A^T u \geq -p\right\},$$

$$(D1)\quad Max\left\{\Theta(x,u) = -x^T Cx - b^T u \mid 2Cx + A^T u = -p,\ u \geq 0\right\}.$$

Anmerkung:

1. Im Problem (*DIII*) sind die Restriktionen in Form von Gleichungen gegeben, und x ist eine freie Variable. Daraus folgt, wenn C eine streng positiv definite Matrix ist (und damit die Inverse C^{-1} existiert), kann man das Problem (*DIII*) durch die Elimination von x vereinfachen. Aus der Beziehung $2Cx + A^T u = -p$ ergibt sich

(1. 11.) $$x = -\frac{1}{2} C^{-1}(A^T u + p).$$

Durch die Substitution von (1. 11.) in die Zielfunktion erhält man:

$$\begin{aligned} x^T Cx + b^T u &= -\frac{1}{4}\left(C^{-1}(A^T u + p)\right)^T \left(C^{-1}(A^T u + p)\right) + b^T u \\ &= \frac{1}{4}\left(A^T u + p\right)^T C^{-1}\left(A^T u + p\right) + b^T u \\ &= \frac{1}{4}\left(A^T u\right)^T C^{-1}\left(A^T u\right) + \frac{1}{4} 2AC^{-1} pu + b^T u + \frac{1}{4} p^T C^{-1} p \\ &= \frac{1}{4} u^T AC^{-1} A^T u + \left(\frac{1}{2} AC^{-1} p + b\right)^T u + const. \\ &= u^T Gu + h^T u + const., \end{aligned}$$

wobei

(1. 12.) $$G := \frac{1}{4} AC^{-1} A^T$$

(1. 13.) $$h := \frac{1}{2} AC^{-1} p + b.$$

Da die Konstante in der Zielfunktion keinen Einfluss auf die Lösung hat, kann man das Problem (*DIII*) wie folgt formulieren:

$$Max\left\{\Theta(u) = -u^T Gu - h^T u \mid u \geq 0\right\}$$

bzw.

(*D*III*) $$Min\left\{\Theta(u) = u^T Guh^T u \mid u \geq 0\right\}.$$

Zur Vervollständigung führen wir die Kuhn-Tucker-Bedingungen für das Problem (*D*III*) an:

(1. 14.) $$\begin{cases} 2Gu - y = -h \\ u \geq 0,\ y \geq 0 \\ u^T y = 0 \end{cases}$$

2. Die Probleme (*DI*), (*DII*) und (*DIII*) bzw. (*D*III*) sind Aufgaben der quadratischen Optimierung.

Lemma 1:
Es sei ein Paar dualer quadratischer Programme (z.B. (*PI*) und (*DI*),…) gegeben. Ist x^1 eine zulässige Lösung des Primalproblems und (x^2, u^2) eine zulässige Lösung des Dualproblems, so gilt

(1. 15.) $$x^{1^T} C x^1 + p^T x^1 \geq -x^{2^T} C x^2 - b^T u^2.$$

Folgerung 1:
Ist das Primal- und auch das Dualproblem zulässig, so haben beide eine optimale Lösung.

Folgerung 2:
Ist bei einem Problem die Zielfunktion über dem zulässigen Bereich nicht beschränkt, so bildet der zulässige Bereich des anderen Problems eine leere Menge.

Satz 3:
Es sei ein Paar dualer Probleme der quadratischen Optimierung gegeben.

a) Stellt x^* eine optimale Lösung des Primalproblems dar, so existiert ein Vektor u^* derart, dass (x^*, u^*) eine optimale Lösung des Dualproblems ist.
b) Stellt (x^*, u^*) eine optimale Lösung des Dualproblems dar, so liefert x^* eine optimale Lösung des Primalproblems.
In beiden Fällen gilt:

(1. 16.) $$Min(x^T C x + p^T x) = Max\left(-x^T C x - b^T u\right).$$

Satz 4:
Es sei ein Paar dualer Probleme der quadratischen Optimierung gegeben.

a) Hat *ein* Problem eine optimale Lösung, so existiert auch eine optimale Lösung des zweiten Problems.
b) Bilden die zulässigen Lösungen eines Problems eine leere Menge, so ist beim zweiten Problem entweder der zulässige Bereich leer, oder die Zielfunktion ist über dem zulässigen Bereich nicht beschränkt