凸优化习题讲义

2021 年 6 月 3 日

Ex 1 Let $C \subset \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \left\{ x \in \mathbb{R}^n | x^\top A x + b^\top x + c \leqslant 0 \right\},$$

with $A \in \mathbf{S}^n, b \in \mathbb{R}^n$, and $c \in \mathbb{R}$.

(a) Show that C is convex if $A \succeq 0$.

(b) Show that the intersection of C and the hyperplane defined by $g^\top x + h = 0$ (where $g \neq 0$) is convex if $A + \lambda g g^\top \succeq 0$ for some $\lambda \in \mathbb{R}$.

证明. (a) 令$x, y \in C$，$\theta \in [0,1]$则

$$\begin{aligned}
&(\theta x + (1-\theta)y)^\top A(\theta x + (1-\theta)y) + b^\top(\theta x + (1-\theta)y) + c \\
=&\theta\left(x^\top A x + b^\top x + c\right) + (1-\theta)(y^\top A y + b^\top y + c) \\
&+ (\theta x + (1-\theta)y)^\top A(\theta x + (1-\theta)y) - \theta x^\top A x - (1-\theta) y^\top A y
\end{aligned}$$

只需要证明$(\theta x + (1-\theta)y)^\top A(\theta x + (1-\theta)y) - \theta x^\top A x - (1-\theta)y^\top A y \leqslant 0$。

$$\begin{aligned}
&(\theta x + (1-\theta)y)^\top A(\theta x + (1-\theta)y) - \theta x^\top A x - (1-\theta) y^\top A y \\
=&(\theta^2 - \theta)(x^\top A x + y^\top A y - 2x^\top A y) \\
=&(\theta^2 - \theta)\left(\sqrt{A}x - \sqrt{A}y\right)^\top \left(\sqrt{A}x - \sqrt{A}y\right)
\end{aligned}$$

又因为当$\theta \in [0,1]$时，$\theta^2 - \theta \leqslant 0$，故上式$\leqslant 0$。

(b) 假设$x, y \in C$，且$g^\top x + h = 0, g^\top y + h = 0$，及$\theta \in [0,1]$。显然，

$$g^\top(\theta x + (1-\theta)y) + h = \theta(g^\top x + h) + (1-\theta)(g^\top y + h) = 0.$$

至于证明$(\theta x+(1-\theta)y)^\top A(\theta x+(1-\theta)y)+b^\top(\theta x+(1-\theta)y)+c\leqslant 0$，由上一问可知，只需要证明

$$x^\top Ax+y^\top Ay-2x^\top Ay\geqslant 0$$

注意到$h^2=(-h)(-h)=(g^\top x)(g^\top x)=(g^\top y)(g^\top y)=(g^\top x)(g^\top y)$，故令$S=\sqrt{A+\lambda gg^\top}$，我们有

$$\begin{aligned}
&x^\top Ax+y^\top Ay-2x^\top Ay\\
=&x^\top Ax+y^\top Ay-2x^\top Ay+\lambda h^2+\lambda h^2-2\lambda h^2\\
=&x^\top Ax+y^\top Ay-2x^\top Ay+\lambda x^\top gg^\top x+\lambda y^\top gg^\top y-2\lambda x^\top gg^\top y\\
=&x^\top(A+\lambda gg^\top)x+y^\top(A+\lambda gg^\top)y-2x^\top(A+\lambda gg^\top)y\\
=&(Sx-Sy)^\top(Sx-Sy)\geqslant 0
\end{aligned}$$

□

Ex 2 Let $\lambda_1(X)\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_n(X)$ denote the eigenvalues of a matrix $X\in\mathbf{S}^n$. Prove that the maximum eigenvalue $\lambda_1(X)$ is convex. Moreover, show that $\sum_{i=1}^k\lambda_i(X)$ is convex on $\mathbf{S}^n$. （请证明过程中务必证明hint）

证明. 令$X=O\Lambda O^\top$为X的特征值分解，其中$O=[o_1,\ldots,o_n]$为正交阵，

$$\Lambda=diag(\lambda_1(X),\lambda_2(X),\ldots,\lambda_n(X))$$

为对角阵。注意到取$V\in\mathbb{R}^{n\times k}$为

$$V=\left[\begin{array}{llll} o_1 & o_2 & \ldots & o_k\end{array}\right]$$

则$tr(V^\top XV)=\sum_{i=1}^k\lambda_i(X)$，且$V^\top V=I$。另一方面，对任意$V\in\mathbb{R}^{n\times k}$使得$V^\top V=I$，记$V=[v_1,\ldots,v_k]$，

$$tr\left(V^\top XV\right)=\sum_{j=1}^n\sum_{i=1}^k\lambda_j(X)\left(v_i^\top o_j\right)^2=\sum_{j=1}^n\lambda_j(X)\sum_{i=1}^k\left(v_i^\top o_j\right)^2$$

注意到，$v_1,\ldots,v_k$可以通过添加另外$n-k$个列向量$v_{k+1},\ldots,v_n$，成为$\mathbb{R}^n$上的一组标准正交基，此时

$$\sum_{i=1}^n\left(v_i^\top o_j\right)^2=1\quad\Rightarrow\quad\sum_{i=1}^k\left(v_i^\top o_j\right)^2\leqslant 1.$$

另一方面，

$$\sum_{j=1}^{n}\sum_{i=1}^{k}\left(v_i^\top o_j\right)^2 = k.$$

因此

$$\begin{aligned} tr\left(V^\top XV\right) \leqslant \max \quad & \sum_{j=1}^{n} a_j\lambda_j(X) \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_j = k \\ & a_j \in [0,1] \end{aligned}$$

RHS这个简单的优化问题，其最大值为$\sum_{i=1}^{k}\lambda_i(X)$。

综上，$\sum_{i=1}^{k}\lambda_i(X) = \max\{tr\left(V^\top XV\right) : V \in \mathbb{R}^{n\times k}, V^\top V = I\}$。由于$tr\left(V^\top XV\right)$是关于$X$的线性函数，$\sum_{i=1}^{k}\lambda_i(X)$是一族关于$X$的线性函数的上确界，因此是关于$X$的凸函数。

□

Ex 3 Find the dual function of the LP

$$\begin{aligned} \min \ & c^\top x \\ \text{s.t.} \ & Gx \leqslant h \\ & Ax = b. \end{aligned}$$

Give the dual problem, and make the implicit equality constraints explicit.

解. Lagrange函数为

$$L(x,\lambda,\mu) = c^\top x + \lambda^\top(Gx-h) + \mu^\top(Ax-b)$$

其中对偶变量$\lambda \geqslant 0$。其对偶函数为

$$\begin{aligned} G(\lambda,\mu) &= \min_x c^\top x + \lambda^\top(Gx-h) + \mu^\top(Ax-b) \\ &= \min_x (c + G^\top\lambda + A^\top\mu)^\top x - h^\top\lambda - b^\top\mu \\ &= \begin{cases} -h^\top\lambda - b^\top\mu & \text{if } c + G^\top\lambda + A^\top\mu = 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

因此对偶问题为

$$\begin{aligned}\max\ \ & -h^\top\lambda - b^\top\mu \\ \text{s.t.}\ \ & c + G^\top\lambda + A^\top\mu = 0 \\ & \lambda \geqslant 0.\end{aligned}$$

□

Ex 4 Consider the equality constrained least-squares problem

$$\begin{aligned}\min\ \ & \|Ax-b\|_2^2 \\ \text{s.t.}\ \ & Gx = h\end{aligned}$$

where $A \in \mathbb{R}^{m\times n}$ with $\mathbf{rank}A = n$, and $G \in \mathbb{R}^{p\times n}$ with $\mathbf{rank}G = p$. Give the KKT conditions, and derive expressions for the primal solution x^* and the dual solution ν^*.

解. Lagrange函数

$$\begin{aligned}L(x,\mu) &= \|Ax-b\|_2^2 + \mu^\top(Gx-h) \\ \nabla_x L(x,\mu) &= 2A^\top(Ax-b) + G^\top\mu\end{aligned}$$

KKT条件为

$$\begin{cases}2A^\top(Ax-b) + G^\top\mu = 0 \\ Gx = h\end{cases}$$

由$\mathbf{rank}A = n$，$A^\top A$可逆，故

$$\begin{aligned}2A^\top Ax =& 2A^\top b - G^\top\mu \\ x =& \left(A^\top A\right)^{-1}\left(A^\top b - \frac{G^\top\mu}{2}\right)\end{aligned}$$

将该式代入到KKT的第二式中，得

$$\begin{aligned}G\left(A^\top A\right)^{-1}\left(A^\top b - \frac{G^\top\mu}{2}\right) =& h \\ G\left(A^\top A\right)^{-1}G^\top\mu =& 2G\left(A^\top A\right)^{-1}A^\top b - 2h\end{aligned}$$

再由$\mathbf{rank}G = p$，$G\left(A^\top A\right)^{-1}G^\top$可逆，故

$$\mu = \left[G\left(A^\top A\right)^{-1}G^\top\right]^{-1}\left(2G\left(A^\top A\right)^{-1}A^\top b - 2h\right)$$

$$x = \left(A^\top A\right)^{-1}\left(A^\top b - G^\top\left[G\left(A^\top A\right)^{-1}G^\top\right]^{-1}\left(G\left(A^\top A\right)^{-1}A^\top b - h\right)\right)$$

□

Ex 5 Suppose $Q \succeq 0$. The problem

$$\begin{aligned} \min\ & f(x) + (Ax-b)^\top Q(Ax-b) \\ \text{s.t.}\ & Ax = b \end{aligned}$$

is equaivalent to the primal equality constrained optimization problem. What is the Newton step for this problem? Is it the same sa that for the primal problem?

解. KKT条件为

$$\begin{cases} \nabla f(x) + 2A^\top Q(Ax-b) + A^\top \mu = 0 \\ Ax - b = 0 \end{cases}$$

牛顿迭代步δ满足

$$\begin{bmatrix} \nabla^2 f(x) + 2A^\top QA & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}.$$

注意到该方程组的第二行$A\delta = 0$，这意味着该方程组的解δ满足

$$\left(\nabla^2 f(x) + 2A^\top QA\right)\delta = \nabla^2 f(x)\delta,$$

即δ满足

$$\begin{bmatrix} \nabla^2 f(x) & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}.$$

即与原问题的牛顿步相同。 □

Ex 6 Suppose we use the infeasible start Newton method to minimize $f(x)$ subject to $a_i^\top x = b_i$, $i = 1, \ldots, p$.

(a) Suppose the initial point $x^{(0)}$ satisfies the linear equality $a_i^\top x^{(0)} = b_i$. Show that the linear equality will remain satisfied for future iterates, i.e., $a_i^\top x^{(k)} = b_i$ for all k.

(b) Suppose that one of the equality constraints becomes satisfied at iteration k, i.e., we have $a_i^\top x^{(k-1)} \neq b_i$, $a_i^\top x^{(k)} = b_i$. Show that at iteration k, *all* the equality constraints are satisfied.

证明. (a) 因为对任意k，其牛顿步$\delta^{(k)}$满足

$$\begin{bmatrix} \nabla^2 f(x^{(k)}) & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} \delta^{(k)} \\ w \end{bmatrix} = -\begin{bmatrix} \nabla f(x^{(k)}) \\ Ax^{(k)} - b \end{bmatrix}.$$

其中$A=[a_1,a_2,\dots,a_p]^\top$. 由$a_i^\top x^{(k)}-b_i=0$可推出$a_i^\top\delta^{(k)}=0$，即$a_i^\top(x^{(k)}+\alpha\delta^{(k)})-b_i=0$对任意的实数$\alpha$成立，因此$a_i^\top x^{(k+1)}-b_i=0$。由题设，$k=0$时，$a_i^\top x^{(0)}-b_i=0$，故对任意$k\geqslant 0$均有$a_i^\top x^{(k)}-b_i=0$。

(b) 考虑第$k-1$步的更新量$\delta^{(k-1)}$，满足

$$\begin{bmatrix} \nabla^2 f(x^{(k-1)}) & A^\top \\ A & 0 \end{bmatrix}\begin{bmatrix} \delta^{(k-1)} \\ w \end{bmatrix} = -\begin{bmatrix} \nabla f(x^{(k-1)}) \\ Ax^{(k-1)}-b \end{bmatrix}.$$

因此必有$A(x^{(k-1)}+\delta^{(k-1)})-b=0$。

断言：$x^{(k)}=x^{(k-1)}+\delta^{(k-1)}$。

反证。如果$x^{(k)}=x^{(k-1)}+\alpha\delta^{(k-1)}$，其中$\alpha\neq 1$，则

$$\left.\begin{aligned} a_i^\top(x^{(k-1)}+\alpha\delta^{(k-1)}) =& b_i \\ a_i^\top(x^{(k-1)}+\delta^{(k-1)}) =& b_i \end{aligned}\right\} \Rightarrow (\alpha-1)a_i^\top\delta^{(k-1)}=0$$

这会得到$a_i^\top\delta^{(k-1)}=0$且$a_i^\top x^{(k-1)}-b_i=0$的结论，与题设矛盾。

因此$Ax^{(k)}=A(x^{(k-1)}+\delta^{(k-1)})=b$，所有等式约束均被满足。

□

Ex 7 Suppose we add the constraint $x^\top x\leqslant R^2$ to the problem (106):

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x)\leqslant 0,\quad i=1,\dots,m \\ & Ax=b \\ & x^\top x\leqslant R^2 \end{aligned}$$

Let $\tilde\phi$ denote the logarithmic barrier function for this modified problem. Find $a>0$ for which $\nabla^2(tf_0(x)+\tilde\phi)\succeq aI$ holds, for all feasible x.

解.

$$\tilde\phi(x)=-\sum_{i=1}^m\log(-f_i(x))-\log(R^2-x^\top x)$$

$$\nabla\tilde\phi(x)=\sum_{i=1}^m\frac{1}{-f_i(x)}\nabla f_i(x)+\frac{1}{R^2-x^\top x}\cdot 2x$$

$$\nabla^2\tilde\phi(x)=\sum_{i=1}^m\frac{1}{f_i(x)^2}\nabla f_i(x)\nabla f_i(x)^\top+\sum_{i=1}^m\frac{1}{-f_i(x)}\nabla^2 f_i(x)+\frac{1}{(R^2-x^\top x)^2}xx^\top+\frac{1}{R^2-x^\top x}I$$

前三项都是半正定的，而且$\nabla^2(tf_0(x))$也是半正定的，所以令$a = \frac{1}{R^2}$即有$\nabla^2(tf_0(x) + \tilde{\phi}) \succeq aI$成立。 $\square$

Ex 8 Consider the problem (106), with central path $x^*(t)$ for $t > 0$, defined as the solution of (111).

For $u > p^*$, let $z^*(u)$ denote the solution of

$$\begin{aligned} \min \quad & -\log(u - f_0(x)) - \sum_{i=1}^m \log(-f_i(x)) \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

Show that the curve defined by $z^*(u)$, for $u > p^*$, is the central path. (In other words, for each $u > p^*$, there is a $t > 0$ for which $x^*(t) = z^*(u)$, and conversely, for each $t > 0$, there is a $u > p^*$ for which $z^*(u) = x^*(t)$).

证明. 对任意$u > p^*$，$z^*(u)$满足

$$\begin{cases} \dfrac{1}{u - f_0(z^*(u))} \nabla f(z^*(u)) + \displaystyle\sum_{i=1}^m \dfrac{1}{-f_i(z^*(u))} \nabla f_i(z^*(u)) + A^\top \nu = 0 \\ Az^*(u) = b \end{cases}$$

令$t = \frac{1}{u - f_0(z^*(u))}$，则$z^*(u)$正好满足$x^*(t)$对应的KKT系统。

反之，对任意一个$t > 0$，$x^*(t)$满足

$$\begin{cases} t\nabla f(x^*(t)) + \displaystyle\sum_{i=1}^m \dfrac{1}{-f_i(x^*(t))} \nabla f_i(x^*(t)) + A^\top \nu = 0 \\ Ax^*(t) = b \end{cases}$$

令$u = \frac{1}{t} + f_0(x^*(t))$，则$x^*(t)$满足$z^*(u)$对应的KKT系统。 $\square$