Seoul National University Optimization Lab.

**Nonlinear programming**

2009 Yun-Hong Min

# Solution of homework VI

## 1 Equality-constrained minimization

**Problem 1.1.** Consider the KKT matrix $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}$ where $P \in \mathbb{S}^n_+$, $A \in \mathbb{R}^{p\times n}$, and $\text{rank}A = p < n$.

(a) Show that each of the following statements is equivalent to nonsingularity of the KKT matrix.

- $\mathcal{N}(P) \cap \mathcal{N}(A) = \{0\}$.
- $Ax = 0, x \neq 0 \ \Rightarrow x^T Px > 0$.
- $F^T PF \succ 0$, where $F \in \mathbb{R}^{n\times(n-p)}$ is a matrix for which $\mathcal{R}(F) = \mathcal{N}(A)$.
- $P + A^T QA \succ 0$ for some $Q \succeq 0$.

(b) Show that if the KKT matrix is nonsingular, then it has exactly $n$ positive and $p$ negative eigenvalues.

**Solution of (a)**

(i) Nonsingularity of KKT matrix $\Rightarrow \mathcal{N}(P) \cap \mathcal{N}(A) = \{0\}$: Assume that KKT matrix is nonsingular and there exists nonzero $x$ such that $Ax = 0$ and $Px = 0$. Then,

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ x \end{bmatrix} = 0$$

which implies KKT matrix is singular, a contradiction.

(ii) $\mathcal{N}(P) \cap \mathcal{N}(A) = \{0\} \ \Leftrightarrow \ Ax = 0, x \neq 0 \ \Rightarrow x^T Px > 0$: Since $\mathcal{N}(P) \cap \mathcal{N}(A) = \{0\}$, $x \notin \mathcal{N}(P)$, and hence $x^T Px \neq 0$. Thus, $x^T Px > 0$ as $P \in \mathbb{S}^n_+$. Conversely, if $x \neq 0$ satisfies $Ax = 0$, then $Px \neq 0$ and $Ax = 0$. Thus, $\mathcal{N}(A) \cap \mathcal{N}(P) = \{0\}$.

(iii) $Ax = 0, x \neq 0 \ \Rightarrow x^T Px > 0 \Leftrightarrow F^T PF \succ 0$, where $F \in \mathbb{R}^{n\times(n-p)}$ is a matrix for which $\mathcal{R}(F) = \mathcal{N}(A)$: If $Ax = 0, x \neq 0$, then $x$ must have the form $x = Fz$, where $z \neq 0$ because $\text{rank}(F) = n - p$. Then we have $x^T Px = z^T F^T PFz > 0$. Conversely, $Ax = 0, x \neq 0$ implies there exists $z \neq 0$ such that $AFz = 0, Fz \neq 0$. Since $F^T PF \succ 0$ and $z \neq 0.$, $x^T Px = z^T F^T PFz > 0$.

(iv) $Ax = 0, x \neq 0 \Rightarrow x^T P x > 0 \Leftrightarrow P + A^T Q A \succ 0$ for some $Q \succeq 0$: Suppose $Ax = 0, x \neq 0$ implies $x^T P x > 0$. Then,

$$x^T(P + A^T A)x = x^T P x + \|Ax\|_2^2 > 0$$

for all nonzero $x$. Thus, $P + A^T Q A \succ 0$ for $Q = I$. Conversely, $P + (A^T Q A) \succ 0$ for some $Q \succeq 0$. Then,

$$x^T(P + A^T Q A)x = x^T P x + x^T A^T Q A x > 0$$

This implies that $x^T P x > 0$ whenever $Ax = 0$.

(v) $\mathcal{N}(P) \cap \mathcal{N}(A) = \{0\} \Rightarrow$ Nonsingularity of KKT matrix: Suppose that KKT matrix is singular, i.e., there are $x, y$, not both zero, such that

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

This means that $Px + A^T z = 0$ and $Ax = 0$, and hence $x^T P x + x^T A^T z = 0$. Since $Ax = 0$, this is equivalent to $x^T P x = 0$, so we have $Px = 0$. Thus, by the assumption, $x = 0$ and $z \neq 0$. But, $A^T z = 0$ for nonzero $z$ contradicts that $A$ is a full-row rank.

**Problem 1.2.** The Euclidean projection of the negative gradient $-\nabla f(x)$ on $\mathcal{N}(A)$ is given by

$$\Delta x_{pg} = \operatorname{argmin}_{Au=0} \| - \nabla f(x) - u\|_2.$$

(a) Let $(v, w)$ be the unique solution of

$$\begin{bmatrix} I & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}.$$

Show that $v = \Delta x_{pg}$.

(b) What is the relation between the projected negative gradient $\Delta x_{pg}$ and the negative gradient of the reduced problem (10.5), assuming $F^T F = I$?

(c) The projected gradient method for solving an equality constrained minimization problem uses the step $\Delta x_{pg}$, and a backtracking line search on $f$. Use the results of part (b) to give some conditions under which the projected gradient method converges to the optimal solution, when started from a point $x^{(0)} \in \mathrm{dom} f$ with $Ax^{(0)} = b$.

**Solution.**

(a) $\Delta x_{pg}$ is a minimizer of equality-constrained minimization problem $\min\{\| - \nabla f(x) - u\|_2^2 | Au = 0\}$. Thus, by KKT condition, there exists $w$ such that

$$I\Delta x_{pg} + A^T w = -\nabla f(x).$$

Moreover, since $\Delta x_{pg}$ is feasible for equality-constrained minimization problem, $A\Delta x_{pg} = 0$.

(b)

(c)

**Problem 1.3.** Show that the reduced objective function $\tilde{f}(z) = f(Fz + \hat{x})$ is strongly convex, and that its Hessian is Lipschitz continuous. Express the strong convexity and Lipschitz constants of $\tilde{f}$ in terms of $K$, $M$, $L$, and the maximum and minimum singular values of $F$.

**Solution.** We make the following assumptions:

(i) $S = \{x | x \in \mathrm{dom} f, f(x) \leq f(x^{(0)}),\ Ax = b\}$ is closed.

(ii) $\nabla^2 f(x) \preceq MI$, and

$$\left\| \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}^{-1} \right\|_2 \leq K.$$

(iii) For $x, \tilde{x} \in S$, $\|\nabla^2 f(x) - \nabla^2 f(\tilde{x})\|_2 \leq L \|x - \tilde{x}\|_2$.

First, we will identify strong convexity of $\tilde{f}$. Suppose $F = U\Sigma V^T$ is SVD of $F$. Then, $F^T F = V(\Sigma^T \Sigma)V^T$. Thus,

$$\nabla^2 \tilde{f}(z) = F^T \nabla^2 f(Fz + \hat{x})F \preceq MF^T F \preceq M(\sigma^*)^2 I$$

where $\sigma^* = \max\{|\sigma_{\max}|, |\sigma_{\min}|\}$ is the largest absolute value of singular values of $F$. Secondly, we will identify Lipschitz constants of $\tilde{f}$.

$$\begin{aligned} \|\nabla^2 \tilde{f}(z_1) - \nabla^2 \tilde{f}(z_2)\|_2 &= \|F^T(\nabla^2 f(Fz_1 + \hat{x}) - \nabla^2 f(Fz_2 + \hat{x}))F\|_2 \\ &\leq L\|F^T\|_2 \|F\|_2 \|F(z_1 - z_2)\|_2 \\ &\leq L\|F\|_2^{5/2} \|z_1 - z_2\|_2 \\ &= L(\sigma^*)^{5/2} \|z_1 - z_2\|_2. \end{aligned}$$

**Problem 1.4.** Show that (10.13) holds, i.e.,

$$f(x) - \inf\{\hat{f}(x + v) | A(x + v) = b\} = \lambda(x)^2/2.$$

**Solution** Suppose $v^*$ is a minimizer of $\inf\{\hat{f}(x + v) | A(x + v) = b\}$. Then, by the KKT conditions,

$$\nabla^2 f(x)v^* + A^T w = -\nabla f(x),\ Av^* = 0.$$

Thus,

$$\begin{aligned} f(x) - \hat{f}(x + v^*) &= -\nabla f(x)^T v^* - (1/2)(v^*)^T \nabla^2 f(x) v^* \\ &= (v^*)^T \nabla^2 f(x) v^* + v^* A^T w - (1/2)(v^*)^T \nabla^2 f(x) v^* \\ &= (1/2)(v^*)^T \nabla^2 f(x) v^* = (1/2)\lambda(x)^2. \end{aligned}$$