Optimality conditions

Optimization Lab.

IE department
Seoul National University

16th November 2009

- Recall our optimization, $\min\{f_0(x)|\ f_i(x) \leq 0, i = 1, \cdots, m,$ $h_j(x) = 0, j = 1, \ldots, p\}$ and $\mathcal{D} = \bigcap_{i=0}^{n} \operatorname{dom} f_i \cap \bigcap_{i=1}^{p} \operatorname{dom} h_i$.
- For convenience, in this chapter, we assume $\mathcal{D} = \mathbb{R}^n$. This may not be very strong assumption especially when feasible region is included in $\mathcal{D}$.

Definition

Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \nu_j h_j(x),$$

where $\lambda \geq 0$ and ν are parameters called *Lagrange multipliers* or *dual variables*.

An example

Recall, for convex optimization $\min\{f(x)|Ax=b\}$, x is optimal iff $\exists\lambda$ s.t. $A^T\lambda = \nabla f(x)$, where $\lambda \in \mathbb{R}^p$. The proof essentially shows the condition is also necessary for a local minimum. We consider an alternate proof of the necessity.

Proof: Assume A is of full row rank. By reordering columns, if necessary, $A = [B, N]$ where $B \in \mathbb{R}^{m\times m}$ is of full column rank. Accordingly, partition $x = \binom{x_B}{x_N}$ so that $\min\{f(x)|Bx_B + Nx_N = b\}$, or, by substituting $x_B = B^{-1}(b - Nx_N)$,

$$\min\left\{g(x_N) := f\left(B^{-1}(b - Nx_N), x_N\right) \middle| x_N \in \mathbb{R}^{n-m}\right\}. \tag{1}$$

An example(*cont'd*)

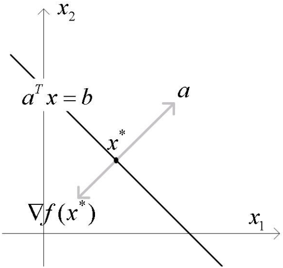
Hence for any local minimum, $x^* = \begin{pmatrix} x_B^* \\ x_N^* \end{pmatrix}$, x_N^* is a local minimum of (1). Thus $\nabla g(x_N^*) = 0$:

$$-N^T B^{-T} \nabla_{x_B} f\left(B^{-1}(b - Nx_N^*), x_N^*\right) + \nabla_{x_N} f\left(B^{-1}(b - Nx_N^*), x_N^*\right) = 0.$$

Letting, $\lambda^* = -B^{-T} \nabla_{x_B} f\left(B^{-1}(b - Nx_N^*), x_N^*\right)$, and using $x_B^* = B^{-1}(b - Nx_N^*)$ we get $\nabla f(x^*) + [B; N]^T \lambda^* = 0$ as desired. $\square$

An example(*cont'd*)

Notice analogy between $\nabla f(x^*) + A^T \lambda^* = 0$ (or, $\nabla f(x^*) \in \mathcal{N}(A)$) and $\nabla f(x^*) = 0$, the necessary condition of unconstrained case. For constrained case, it suffices that the gradient $\nabla f(x)$ at $x = x^*$ vanishes along every direction into subspace $Ax = b$ instead of $\mathbb{R}^n$.



An example(*cont'd*)

Now we consider the second order necessary condition of local minimum of an unconstrained case: Hessian of objective at a local minimum is PSD (See Exercise). Applying this to (1), if x^* is a local minimum then $\nabla^2 g(x_N^*) \succeq 0$.

An example(*cont'd*)

But,

$$\begin{aligned}
\nabla^2 g(x_N) &= \nabla\Big(-N^T B^{-T} \nabla_{x_B} f\big(B^{-1}(b - Nx_N), x_N\big) + \nabla_{x_N} f\big(B^{-1}(b - Nx_N), x_N\big)\Big) \\
&= -N^T B^{-T} \left[\nabla^2_{x_B x_B} f(B^{-1}(b - Nx_N), x_N); \nabla^2_{x_B x_N} f(B^{-1}(b - Nx_N), x_N)\right] \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} \\
&\quad + \left[\nabla^2_{x_N x_B} f(B^{-1}(b - Nx_N), x_N); \nabla^2_{x_N x_N} f(B^{-1}(b - Nx_N), x_N)\right] \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} \\
&= N^T B^{-T} \nabla^2_{x_B x_B} f(B^{-1}(b - Nx_N), x_N) B^{-1} N - N^T B^{-T} \nabla^2_{x_B x_N} f(B^{-1}(b - Nx_N), x_N) \\
&\quad - \nabla^2_{x_N x_B} f(B^{-1}(b - Nx_N), x_N) B^{-1} N + \nabla^2_{x_N x_N} f(B^{-1}(b - Nx_N), x_N),
\end{aligned}$$

where,

$$\nabla^2 f(x^*) = \begin{bmatrix} \nabla^2_{x_B x_B} f(x^*) & \nabla^2_{x_B x_N} f(x^*) \\ \nabla^2_{x_N x_B} f(x^*) & \nabla^2_{x_N x_N} f(x^*) \end{bmatrix}.$$

An example(*cont'd*)

Thus

$$\begin{aligned}\nabla^2 g(x_N^*) &= N^T B^{-T} \nabla^2_{x_B x_B} f(x^*) B^{-1} N - N^T B^{-T} \nabla^2_{x_B x_N} f(x^*) \\ &\quad - \nabla^2_{x_N x_B} f(x^*) B^{-1} N + \nabla^2_{x_N x_N} f(x^*) \\ &= \begin{bmatrix} -N^T B^{-T} & I \end{bmatrix} \begin{bmatrix} \nabla^2_{x_B x_B} f(x^*) & \nabla^2_{x_B x_N} f(x^*) \\ \nabla^2_{x_N x_B} f(x^*) & \nabla^2_{x_N x_N} f(x^*) \end{bmatrix} \begin{bmatrix} -B^{-1} N \\ I \end{bmatrix}.\end{aligned}$$

From positive semidefiniteness of $\nabla^2 g(x_N^*)$,

$$\forall y_B \in \mathbb{R}^{n-m},\ y_N^T \nabla^2 g(x_N^*) y_N = (-y_N^T N^T B^{-T}, y_N^T) \nabla^2 f(x^*) \begin{pmatrix} -B^{-1} N y_N \\ y_N \end{pmatrix} \geq 0.$$

Note that $Ay = [B; N] \binom{y_B}{y_N} = 0$ iff $y_B = -B^{-1} N y_N$. Thus $\nabla^2 f(x^*) = \nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla^2 h_i(x^*)$ is PSD on $\mathcal{N}(A)$.

First order necessary condition

The first and second order necessary conditions for $\min\{f(x)|Ax = b\}$ can be generalized for general equality constrained cases.

Theorem

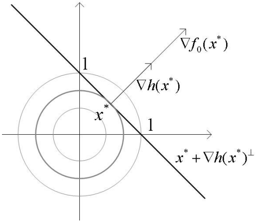
Consider equality constrained optimization $\min\{f(x)|\ h_i(x) = 0, i = 1, \cdots, p\}$. *Assume* x^* *is a local minimum and* $\nabla h_1(x^*)$, ..., $\nabla h_p(x^*)$ *are linearly indep. Then there is unique* λ^* *such that*

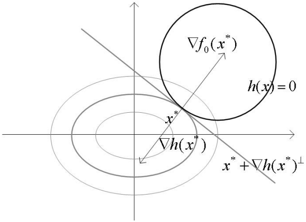
$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) = 0. \tag{2}$$

First order necessary condition(*cont'd*)

Example

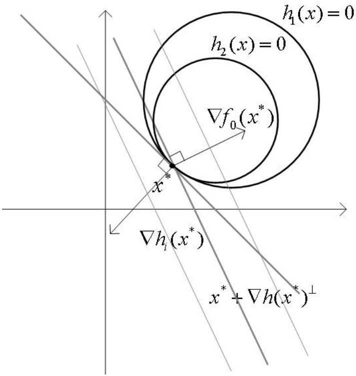
$$\min x_1^2 + x_2^2$$
$$\text{s.t. } x_1 + x_2 = 1.$$





First order necessary condition(*cont'd*)

- Without regularity assumption, λ^* may not exist.



First order necessary condition(*cont'd*)

Proof

By using Implicit Function Theorem, can be proven similarly with the linear constrained case in which $x_B = B^{-1}(b - Nx_N)$ was an implicit function. But, here we use penalty approach: Let x^* be a local min of $\min\{f(x) : h(x) = 0\}$ and for $k \in \mathbb{N}$, consider the objective augmented with penalties,

$$P_k(x) = f(x) + \frac{k}{2}\|h(x)\|_2^2 + \frac{\alpha}{2}\|x - x^*\|_2^2,$$

where α is a positive scalar.

First order necessary condition(*cont'd*)

Proof(*cont'd*)

Let x^k be optimal solution of

$$\min P_k(x)$$
$$\text{s.t } x \in S := \{x | \|x - x^*\| \leq \epsilon\}$$

Notice that for all k,

$$P_k(x^k) = f(x^k) + \frac{k}{2}\|h(x^k)\|_2^2 + \frac{\alpha}{2}\|x^k - x^*\|_2^2 \leq P_k(x^*) = f(x^*). \tag{3}$$

As $f(x^k)$ is bounded below on S, (3) implies $h(x^k) \to 0$ as $k \to \infty$.

First order necessary condition(*cont'd*)

Proof(*cont'd*)

Also (3) implies $f(x^k) + \frac{\alpha}{2}\|x^k - x^*\|_2^2 \leq f(x^*)$. Hence for every limit point of $\bar{x}$ of $\{x^k\}$, we get

$$f(\bar{x}) + \frac{\alpha}{2}\|\bar{x} - x^*\|_2^2 \leq f(x^*). \tag{4}$$

But since $h(x^k) \to h(\bar{x}) = 0$, $\bar{x}$ is feasible and we have $f(\bar{x}) \geq f(x^*)$.
Combining this with (4) we get $\bar{x} - x^* = 0$ or $\bar{x} = x^*$.
We have seen that $\{x^k\}$ converges to x^* and hence for sufficiently large k's, x^k is in the interior of S and x^k is a minimum of the unconstrained problem $\min P_k(x)$. Therefore, for such k's

$$0 = \nabla P_k(x^k) = \nabla f(x^k) + kDh(x^k)^T h(x^k) + \alpha(x^k - x^*). \tag{5}$$

First order necessary condition(*cont'd*)

Proof(*cont'd*)

Also for sufficiently large k's, $Dh(x^k)$ is of full row rank as $Dh(x^*)$. Thus $Dh(x^k)Dh(x^k)^T$ is invertible. Multiplying (5) with $(Dh(x^k)Dh(x^k)^T)^{-1}Dh(x^k)$, we get

$$kh(x^k) = -\left(Dh(x^k)Dh(x^k)^T\right)^{-1} Dh(x^k)\left(\nabla f(x^k) + \alpha(x^k - x^*)\right).$$

Hence, by taking limit as $k \to \infty$, $kh(x^k)$ converges to

$$\lambda^* := -\left(Dh(x^*)Dh(x^*)^T\right)^{-1} Dh(x^*)\nabla f(x^k).$$

Thus taking limit as $k \to \infty$ in (5), we get

$$\nabla f(x^*) + Dh(x^*)^T\lambda^* = 0. \quad \square$$

Second order necessary condition

Theorem

Assume x^* is a local minimum , $\nabla h_1(x^*)$, ..., $\nabla h_p(x^*)$ are linearly indep, and f and h_i are twice differ'ble. Assume λ^* be the unique multiplier satisfying the first order necessary condition. Then

$$z^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla^2 h_i(x^*) \right) z \geq 0, \forall\, z : \nabla h_i(x^*)^T z = 0. \tag{6}$$

In other words, $\nabla_x^2 L(x^*, \lambda^*)$ is PSD on the nullspace of $Dh(x^*)$.

Second order necessary condition(*cont'd*)

Proof

Since x^k is an unconstrained minimum of $\min P_k(x)$, the following matrix

$$\nabla^2 P_k(x^k) = \nabla^2 f(x^k) + k\sum_{i=1}^{m} h_i(x^k)\nabla^2 h_i(x^k) + kDh(x^k)^T Dh(x^k) + \alpha I, \quad (7)$$

is positive semidefinite for all suff large k and $\alpha > 0$.
Take any z such that $Dh(x^*)z = 0$ and let z^k be the projection of z onto the nullspace of $Dh(x^k)$:

$$z^k = z - Dh(x^k)^T \left(Dh(x^k)Dh(x^k)^T\right)^{-1} Dh(x^k)z. \quad (8)$$

Second order necessary condition(*cont'd*)

Proof(*cont'd*)

Since $Dh(x^k)z^k = 0$, and $\nabla^2 P_k(x^k)$ is PSD, we have

$$0 \leq (z^k)^T \nabla^2 P_k(x^k) z^k = (z^k)^T \left(\nabla^2 f(x^k) + k \sum_{i=1}^{m} h_i(x^k) \nabla^2 h_i(x^k) \right) z^k + \alpha \|z^k\|^2.$$

Since $kh_i(x^k) \to \lambda_i^*$, and from (8) together with $x^k \to x^*$ and $Dh(x^*)z = 0$, we have $z^k \to z$, we get

$$0 \leq z^T \left(\nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 h_i(x^*) \right) z + \alpha \|z\|^2.$$

By taking α arbitrarily close to 0, we get the second order condition. $\square$

Second order sufficient condition

Theorem

Consider minimization $\min\{f(x) : h(x) = 0\}$ *where* $h : \mathbb{R}^n \to \mathbb{R}^p$, f *and* h_i *are twice differ'ble. Assume* $x^* \in \mathbb{R}^n$ *and* $\lambda^* \in \mathbb{R}^p$ *satisfy*

$$\nabla_\lambda L(x^*, \lambda^*) = 0, \quad \nabla_x L(x^*, \lambda^*) = 0$$
$$z^T \left(\nabla^2 f_0(x^*) + \sum_{i=1}^p \lambda_i^* \nabla^2 h_i(x^*)\right) z > 0, \forall\, z \neq 0 : \nabla h_i(x^*)^T z = 0.$$

Then x^* *is a strict local minimum of* f*. In fact, there is* $\gamma > 0$, *and* $\epsilon > 0$ *such that*

$$f(x^*) + \frac{\gamma}{2}\|x - x^*\|^2 \leq f(x), \ \forall x : h(x) = 0 \ \mathit{and} \|x - x^*\| < \epsilon.$$

Proof Omitted. □

First order necessary condition

Theorem

Assume x^* *is a local minimum of* $\min\{f_0(x) : f_1(x) \leq 0, \ldots, f_m(x) \leq 0, h_1(x) = 0, \ldots, h_p(x) = 0\}$ *where* f_i *and* h_i *are differ'ble. Assume* x^* *is regular. Then there are unique* $\lambda^* \geq 0$ *and* ν^* *such that*

1. *Dual feasibility:* $\lambda^* \geq 0$,
2. *Complementary slackness:* $\lambda_i^* f_i(x^*) = 0,\ i = 1, \cdots, m$,
3. *Gradient of Lagrangian with respect to* x *vanishes at* x^* *when* $\lambda = \lambda^*$ *and* $\nu = \nu^*$*:*

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \nu_j^* \nabla h_j(x^*) = 0.$$

First order necessary condition
Proof

It suffices to show that $\lambda^* \geq 0$. (Why?) Introduce $f_j^+(x) = \max\{0, f_j(x)\}$ and

$$P_k(x) = f_0(x) + \frac{k}{2}\|h(x)\|_2^2 + \frac{k}{2}\sum_{i=1}^{m}(f_i^+(x))^2 + \frac{\alpha}{2}\|x - x^*\|_2^2,$$

where α is a positive scalar. Notice that $(f_i^+(x))^2$ is differentiable with gradient $2f_i^+(x)\nabla f_i(x)$. A similar argument we used for equality case, the unique multipliers are given by

$$\nu_i^* = \lim_{k\to\infty} kh_i(x^k), \ i = 1, \ldots, p,$$
$$\lambda_i^* = \lim_{k\to\infty} kf_i^+(x^k), \ i = 1, \ldots, m.$$

Since $f_i^+(x^k) \geq 0$, we get $\lambda_i^* \geq 0$. $\square$

The regularity of x^* is quite strong assumption for inequality constraints. It can be replaced by a weaker assumption.

Second order necessary condition

Theorem

Assume x^* *is a local minimum of* $\min\{f_0(x) : f_1(x) \leq 0, \ldots, f_m(x) \leq 0, h_1(x) = 0, \ldots, h_p(x) = 0\}$ *where* f_i *and* h_i *are twice differ'ble. Let* $A(x^*)$ *be the set of active constraints at* x^**. Then the unique Lagrangian multipliers* $\lambda^* \geq 0$ *and* ν^* *satisfying the first order condition also satisfy*

$$z^T \left(\nabla^2 f_0(x^*) + \sum_{i=1}^p \lambda_i^* \nabla^2 f_i(x^*) + \sum_{i=1}^p \nu_i^* \nabla^2 h_i(x^*)\right) z \geq 0,$$
$$\forall\, z : \nabla f_i(x^*)^T z = 0,\ i \in A(x^*) \text{ and } \nabla h_i(x^*)^T z = 0,\ i = 1, \ldots, p.$$

Proof From second order necessary conditions for equality const case. □

Second order sufficient condition

Theorem

Consider minimization $\min\{f_0(x) : f_1(x) \leq 0, \ldots, f_m(x) \leq 0, h_1(x) = 0, \ldots, h_p(x) = 0\}$ *where f_i and h_i are twice differ'ble. Let $x^* \in R^n$, λ^*, and μ^* satisfy*

$$\nabla_\lambda L(x^*, \lambda^*) = 0, \;\; \nabla_x L(x^*, \lambda^*) = 0,$$
$$f_i(x^*) \leq 0, \; i = 1, \ldots, m,$$
$$h_i(x^*) = 0, \; i = 1, \ldots, p,$$
$$\mu_i^* \geq 0, \; i = 1, \ldots, p, \; \mu_i^* = 0, \forall i \in A(x^*),$$
$$z^T \left(\nabla^2 f_0(x^*) + \textstyle\sum_{i=1}^p \lambda_i^* \nabla^2 f_i(x^*) + \sum_{i=1}^p \nu_i^* \nabla^2 h_i(x^*)\right) z > 0,$$
$$\forall\, z \neq 0 : \nabla f_i(x^*)^T z = 0, \; i \in A(x^*) \text{ and } \nabla h_i(x^*)^T z = 0, \; i = 1, \ldots, p.$$

Then x^* is a strict local minimum.

Proof Omitted. □

KKT conditions

The following four conditions are called KKT conditions (for a problem with differentiable f_i and h_i):

1. Primal feasibility: $f_i(x) \le 0$, $i = 1, \cdots, m$; $h_j(x) = 0$, $j = 1, \cdots, p$,
2. Dual feasibility: $\lambda \ge 0$,
3. Complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \cdots, m$,
4. Gradient of Lagrangian with respect to x vanishes:

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{j=1}^{p} \nu_j \nabla h_j(x) = 0.$$

We have seen that under regularity assumption KKT conditions are necessary for a local minimum.

KKT conditions are sufficient for convex optimization.

Proposition

Suppose optimization $\min\{f_0(x) : f_1(x) \le 0, \ldots, f_m(x) \le 0, h_1(x) = 0, \ldots, h_p(x) = 0\}$ *is convex. If* $\tilde{x}$ *and* $(\tilde{\lambda}, \tilde{\nu})$ *satisfy KKT conditions, then they are optimal.*

Proof From convexity and the 4th condition,

$$L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) := f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{x}) + \sum_{j=1}^{p} \tilde{\nu}_j h_j(\tilde{x}) \le L(x, \tilde{\lambda}, \tilde{\nu}), \ \forall x.$$

Since $\tilde{\lambda} \ge 0$,

$$L(x, \tilde{\lambda}, \tilde{\nu}) \le f_0(x), \ \forall \text{ feasible } x.$$

Thus, $L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) \le p^*$. But, from complementary slackness, $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$ and hence, $f_0(\tilde{x}) = p^*$. $\square$

Thus under the regularity assumption KKT conditions are necessary and sufficient for optimality. The regularity can be replaced by a weaker form of constraint qualification such as Slater's condition.

Corollary

Suppose there is feasible solution $\bar{x}$ such that $f_i(\bar{x}) < 0\ \forall i$. Then x is optimal for a convex optimization iff there exist λ, ν satisfying KKT conditions with x.

Examples

Consider equality constrained convex quadratic minimization

$$\begin{array}{ll} \min & (1/2)x^T P x + q^T x + r \\ \text{s.t.} & Ax = b, \end{array}$$

where $P \in \mathbb{S}_+^n$. KKT conditions are $Ax^* = b$, $Px^* + q + A^T \nu^* = 0$, or

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}.$$

Examples(*cont'd*)

Consider following optimization:

$$\begin{array}{ll} \min & -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \text{s.t.} & x \geq 0,\ \mathbf{1}^T x = 1, \end{array}$$

where $\alpha_i > 0$. KKT conditions for this problem are

$$x^* \geq 0,\ \mathbf{1}^T x^* = 1,\ \lambda^* \geq 0,\ \lambda_i^* x_i^* = 0,\ i = 1, \cdots, n,$$
$$-1/(\alpha_i + x_i^*) - \lambda_i^* + \nu^* = 0,\ i = 1, \cdots, n.$$

Solving the equations, we have

$$x_i^* = \begin{cases} 1/\nu^* - \alpha_i, & \nu^* \leq 1/\alpha_i \\ 0 & \nu^* \geq 1/\alpha_i \end{cases},\ \text{or } x_i^* = \max\{0, 1/\nu^* - \alpha_i\}.$$

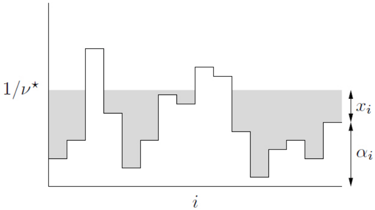
Since $\mathbf{1}^T x^* = 1$, we can obtain

$$\sum_{i=1}^{n} \max\{0, 1/\nu^* - \alpha_i\} = 1.$$

Examples(*cont'd*)

This solution method is called *water-filling* for the following reason:

- α_i is ground level above patch i.
- $1/\nu^*$ is target depth for flood.
- Total amount of water used is $\sum_i \max\{0, 1/\nu^* - \alpha_i\}$.
- We increase flood level until we have used total amount of water equal to one. Then, final depth of water above patch i is x_i^*.



Exercises

1. Use Lagrangian to solve the followings.
(a) min $\{\|x\|^2 : \sum_{i=1}^n x_i = 1\}$.
(b) min $\{\sum_{i=1}^n x_i : \|x\|^2 = 1\}$.
(c) min $\{\|x\|^2 : x^T Qx = 1\}$, where Q is PD.

2. Let x^* be an unconstrained local minimum of $f : \mathbb{R}^n \to \mathbb{R}$. Also assume f is twice differentiable in an open set S. Then $\nabla^2 f(x^*)$ is positive semidefinite.

3. Solve the following problem

$$\min (x-a)^2 + (y-b)^2 + xy$$
$$\text{s.t. } 0 \leq x \leq 1,\ 0 \leq x \leq 1,$$

for all possible values of a and b.

Exercises

4. Consider

$$\min -(x_1x_2 + x_2x_3 + x_3x_1)$$
$$\text{s.t. } x_1 + x_2 + x_3 = 3.$$

Show $x^* = (1,1,1)^T$ is a strict local minimum.

5. Verify the Schwartz inequality, $x^T y \leq \|x\|\|y\|$ by solving the problem $\max\{x^T y : \|x\|^2 = 1, ; \|y\|^2 = 1\}$.
Similarly, for any PD matrix Q, prove

$$(x^T y)^2 \leq (x^T y)^2 \leq (x^T Qx)(y^T Q^{-1} y)$$

by solving $\min\{y^T x : x^T Qx \leq 1\}$.

Exercises

6. Show if the constraints are linear, the regularity assumption is not needed for the second order necessary conditions except that the multipliers are not necessarily unique.

7. Consider convex optimization $\min\{f_0(x) : f_i(x) \leq 0,\ i = 1, \ldots, m\}$. Assume x^* satisfies KKT conditions. Show that

$$\nabla f_0(x^*)^T (x - x^*) \geq 0,$$

for all feasible solution x.