
Solution of homework V

Problem 0.1. Use Lagrangian to solve the followings:

(a) $\min\{\|x\|^2 | \sum_{i=1}^n x_1 = 1\}$: **Solution.** We want to find x^* satisfying sufficient condition for optimality. Since $L(x,\lambda) = \|x\|^2 + \lambda(\sum_{i=1}^n x_i - 1)$, x^* and λ^* should satisfy

$$\nabla_x L = 2x + \lambda \mathbf{1} = 0 \text{ and } \sum_{i=1}^n x_i = 1.$$

Thus, $x^* = (-\lambda^*/2)\mathbf{1}$ and by $\sum_{i=1}^n x_i^* = 1$, $\lambda^* = -2/n$. Thus, $x_i^* = 1/n$ for all $i = 1,\ldots,n$. Moreover, $\nabla^2 L(x^*,\lambda^*) = 2 > 0$. Thus, $x^* = (1/n)\mathbf{1}$ is optimal. □

(b) $\min\{\sum_{i=1}^n x_i | \|x\|^2 = 1\}$: **Solution.** Since $L(x,\lambda) = \sum_{i=1}^n x_i + \lambda(\|x\|^2 - 1)$, x^* and λ^* should satisfy

$$\nabla_x L = \mathbf{1} + 2\lambda^* x^* = 0,\ \|x^*\| = 1, \text{and}$$
$$\nabla^2 L(x^*,\lambda^*) = 2\lambda^* I \succ 0.$$

Thus, $x_i^* = -1/(2\lambda^*)$ for all $i = 1,\ldots,n$, and $\|x^*\|^2 = \sum_{i=1}^n (x_i^*)^2 = n/(4(\lambda^*)^2) = 1$ implies $\lambda^* = \pm\sqrt{n}/2$. Since $\lambda^* > 0$, $x_i^* = -1/\sqrt{n}$ for all $i = 1,\ldots,n$. □

(c) $\min\{\|x\|^2 | x^T Q x = 1\}$, where Q is PD :**Solution.** Skip.

Problem 0.2. 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.

Solution. Suppose $\nabla^2 f(x^*)$ is not positive semi-definite. Then there exists d such that $d^T \nabla^2 f(x^*) d < 0$. Then, since f is twice differentiable we have

$$f(x^* + \lambda d) = f(x^*) + \lambda \nabla f(x^*)^T d + \frac{1}{2} d^T \nabla^2 f(x^*) d + \lambda^2 \|d\|^2 \alpha(x^*, \lambda d)$$

where $\alpha(x^*, \lambda d) \to 0$ as $\lambda \to 0$. By rearranging

$$\frac{f(x^* + \lambda d) - f(x^*)}{\lambda^2} = \frac{1}{2} d^T \nabla^2 f(x^*) d + \|d\|^2 \alpha(\cdots).$$

Since $d^T \nabla^2 f(x^*) d < 0$ and $\alpha(\cdots) \to 0$ as $\lambda \to 0$, $f(x^* + \lambda d) - f(x^*) < 0$ for all $\lambda > 0$ sufficiently small, a contradiction. □

Problem 0.3. Solve the following problem

$$\begin{array}{ll} \min & (x-a)^2 + (y-b)^2 + xy \\ \text{sub.to.} & 0 \le x \le 1, 0 \le y \le 1, \end{array}$$

for all possible values of a and b.

Solution.

Problem 0.4. Consider

$$\begin{array}{ll} \min & -(x_1x_2 + x_2x_3 + x_3x_1) \\ \text{sub. to.} & x_1 + x_2 + x_3 = 3. \end{array}$$

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

Solution. It suffices to show that x^* satisfies second order sufficient conditions for equality constrained case. Consider $L(x,\lambda) = -(x_1x_2 + x_2x_3 + x_3x_1) + \lambda(x_1 + x_2 + x_3 - 3)$. First, $\nabla_x L(x^*, \lambda^*) = 0$ holds when $\lambda^* = 2$ and $x^* = [1,1,1]^T$. Second, it is easy to check $\nabla_\lambda L(x^*, \lambda^*) = 0$. Finally,

$$\nabla^2 L = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

and $z^T \nabla^2 L z = -2(z_1x_2 + z_2z_3 + z_3z_1) = z_2^2 + z_3^2 + (z_2+z_3)^2 > 0$ for all z, $z_1 + z_2 + z_3 = 0$. □

Problem 0.5. Verify the Schwartz inequality, $x^T y \le \|x\|\|y\|$ by solving the problem $\max\{x^T y \| \|x\|^2 = 1, \|y\|^2 = 1\}$. Similarly, for any PD matrix Q, prove

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

by solving $\min\{y^T x | x^T Q x \le 1\}$.

Solution.

(i) By the second-order sufficient condition for the equality constrained problem, the optimum solution of a given optimization problem is $x^* = y^* (\lambda^* = 1/2)$. Thus, for any x and y satisfying $\|x\|_2 = \|y\|_2 = 1$, $x^T y \le 1$. Now, for any u, v

$$u^T v = (\|u\|_2 \frac{u}{\|u\|_2})^T (\|v\|_2 \frac{v}{\|v\|_2}) = \|u\|_2 \|v\|_2 (\frac{u}{\|u\|_2})^T (\frac{v}{\|v\|_2}) \le \|u\|_2 \|v\|_2.$$

(ii) skip.

Problem 0.6. 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.

Solution. When the constraints are linear, $Ax = 0$, the gradient of the constraints is just coefficient vector of the constraints, i.e., A, which is the same for all feasible x. Thus if the row rank of A is not full, by the Gaussian elimination we can choose linearly independent rows B of A. Then, use $Bx = 0$ instead of A which is equivalent to the original problem. But, by choice of linearly independent rows of A, the corresponding multipliers can be changed. □

Problem 0.7. 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.

Solution. By the assumption, x^* and corresponding Lagrangian multiplier λ^* satisfy

$$f_i(x^*) \leq 0, \quad i = 1, \ldots, m \tag{1}$$
$$\lambda^* \geq 0, \tag{2}$$
$$\lambda_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m \tag{3}$$
$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0. \tag{4}$$

Let $I_{x^*} = \{i|f_i(x^*) = 0, i = 1, \ldots, m\}$. Then, for any feasible x, $f_i(x) \leq f_i(x^*) = 0$ for $i \in I_{x^*}$. Thus, for $i \in I_{x^*}$

$$f_i(x^* + \lambda(x - x^*)) = f_i((1 - \lambda)x^* + \lambda x) \leq (1 - \lambda)f_i(x^*) + \lambda f_i(x),$$

and

$$f_i(x^* + \lambda(x - x^*)) - f_i(x^*) \leq \lambda(f_i(x) - f_i(x^*))$$
$$\tfrac{f_i(x^*+\lambda(x-x^*))-f_i(x^*)}{\lambda} \leq f_i(x) - f_i(x^*) \leq 0.$$

Thus,

$$\nabla f_i(x^*)^T(x - x^*) = \lim_{\lambda \to 0} \frac{f_i(x^* + \lambda(x - x^*)) - f_i(x^*)}{\lambda} \leq 0$$

for all feasible x and $i \in I_{x^*}$. Since for $i \notin I_{x^*}$ $\lambda_i^* = 0$,

$$\sum_{i=1}^{m} \lambda_i^* \nabla f(x^*)^T(x - x^*) \leq 0.$$

From (4),

$$\nabla f_0(x^*)^T(x - x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f(x^*)^T(x - x^*) = 0.$$

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