Problem Set 5 Solutions

March 29, 2005

2.2.1

We have

$$\nabla f(x) = \begin{pmatrix} x_1 \\ x_2 \\ .1x_3 + .55 \end{pmatrix}$$

and

$$\nabla^2 f(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since the Hessian is positive definite for all $x \in X$, $f(x)$ is convex over the set X. Thus satisfying the first order necessary condition is sufficient for x^* to be a global minimum. We have

$$\nabla f(x^*)'(x - x^*) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & .55 \end{pmatrix} \begin{pmatrix} x_1 - 1/2 \\ x_2 - 1/2 \\ x_3 - 0 \end{pmatrix} = \frac{1}{2}(x_1 + x_2 + x_3) + 0.05x_3 - \frac{1}{2}.$$

Since $x_1 + x_2 + x_3$ is constrained to equal 1, we have

$$\nabla f(x^*)'(x - x^*) = 0.05x_3 \geq 0, \quad \forall\, x \in X.$$

Thus $x^* = (\frac{1}{2}, \frac{1}{2}, 0)$ is optimal.

2.6.1

Convergence from both points to the optimal value $x^* = (1, 0, 0)$ is very fast (about 4-6 iterations).

3.1.3

The problem is

$$\min f(x) = \|y - x\| + \|z - x\|$$

$$\text{subject to } h(x) = 0.$$

Using the Lagrange Multiplier Theorem, we have that if x^* is optimal and regular, then there exists a scalar λ^* such that

$$\frac{x^* - y}{\|x^* - y\|} + \frac{x^* - z}{\|x^* - z\|} + \lambda^* \nabla h(x^*) = 0.$$

This is equivalent to

$$\lambda^* \nabla h(x^*) = \frac{y - x^*}{||y - x^*||} + \frac{z - x^*}{||z - x^*||},$$

which means that the vector $\nabla h(x^*)$ lies in between the two unit vectors. Then for the sum of the unit vectors to be collinear with $\nabla h(x^*)$, $\nabla h(x^*)$ must bisect the angle formed by $x^* - y$ and $x^* - z$. Thus $\phi_y = \phi_z$.

3.1.5

(a) Let P_k be the problem

$$\begin{aligned} &\min x'Qx \\ &\text{s.t. } \|x\|^2 = 1 \\ &\qquad e_i'x = 0, \quad i = 1, \dots, k-1, \end{aligned}$$

for $1 \le k \le n$. The feasible region of P_k is contained in the feasible region of P_{k-1}. Therefore

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n.$$

(b) By construction, we have $||e_i|| = 1$ for all i, and $e_i'e_k = 0$ for all i and k with $i \neq k$, i.e., the nonzero vectors $e_1, \dots, e_n$ are mutually orthogonal. Therefore $e_1, \dots, e_n$ are linearly independent.

(c) From (b), $e_1, \dots, e_n$ are regular points for $P_1, \dots, P_n$, respectively. By the Lagrange Multiplier Theorem, there exists a unique $\mu^* \in \Re^i$ for each problem P_i such that

$$2Qe_i + 2\mu_i^* e_i + \mu_1^* e_1 + \dots \mu_{i-1}^* e_{i-1} = 0. \tag{1}$$

Pre-multiplying by e_i', we get

$$2 \underbrace{e_i'Qe_i}_{=\lambda_i} + 2\mu_i^* \underbrace{e_i'e_i}_{=1} = 0,$$

which is satisfied by $\mu_i^* = -\lambda_i$. Thus we can view $-\lambda_i$ as the Lagrange multiplier for P_i associated with the constraint $||x||^2 = 1$.

Now consider problem P_1. Equation (1) yields $2Qe_1 + 2\mu_1^* e_1 = 2Qe_1 - 2\lambda_1 e_1 = 0$, and so e_1 is an eigenvector corresponding to the eigenvalue λ_1 of Q. Now assume that $e_1, \dots, e_{i-1}$ are eigenvectors corresponding to the eigenvalues $\lambda_1, \dots, \lambda_{i-1}$ of Q. Then pre-multiplying Eq. (1) by e_j' for $j < i$, we have

$$2e_j'Qe_i + \mu_j^* e_j'e_j = 2e_j\lambda_j e_i - \mu_j^* = -\mu_j^* = 0.$$

Substituting the values $\mu_i^* = -\lambda_i, \mu_1^* = \cdots = \mu_{i-1} = 0$ into Eq. (1), we have

$$2Qe_i = 2\lambda_i e_i,$$

and thus e_i is an eigenvector corresponding to the eigenvalue λ_i of Q. By induction, we have that $\lambda_1, \ldots, \lambda_n$ are eigenvalues of Q, and $e_1, \ldots, e_n$ are corresponding eigenvectors.

3.1.7

For the cost function, we have

$$f(x) = \sum_{j=1}^{m} \|x - a_j\|^2 = m\|x\|^2 + \sum_{j=1}^{m} \|a_j\|^2 - 2x' \sum_{j=1}^{m} a_j,$$

which is equal to

$$m\|x - \frac{1}{m}\sum_{j=1}^{m} a_j\|^2 + \sum_{j=1}^{m} \|a_j\|^2 - \frac{1}{m}\|\sum_{j=1}^{m} a_j\|^2.$$

Hence

$$\min_{||x||^2=1} f(x) = \min_{||x||^2=1} ||x - \hat{a}||^2$$

and

$$\max_{||x||^2=1} f(x) = \max_{||x||^2=1} ||x - \hat{a}||^2,$$

i.e. the two optimization problems are equivalent to finding the points on the unit sphere that are at minimum and maximum distance from the center of the gravity $\hat{a}$. If $\hat{a} \neq 0$, these are the points of intersection of the unit sphere with the line that connects the origin and $\hat{a}$. If $\hat{a} = 0$, then all feasible points have the same cost.

To solve the problem using Lagrange multiplier theorem, we apply the 1st order necessary condition. Since all feasible points are regular, we have that for a local minimum or a local maximum x^* there exists a scalar λ^* such that

$$2\sum_{j=1}^{m}(x^* - a_j) + 2\lambda^* x^* = 0.$$

Assuming that $\hat{a} \neq 0$, we see that $m + \lambda^* \neq 0$, so

$$x^* = \frac{m}{m + \lambda^*}\hat{a}.$$

Thus all local maxima and local minima lie on the line connecting the origin with $\hat{a}$, as well as on the surface of the unit sphere. There are exactly two such points. Since the constraint set is compact there exists a global minimum and a global maximum, so one of the two points is the global maximum and the other is the global minimum.

3.1.11

Let x^* be a regular point and a local minimum of the problem

$$\min_{h(x)=0} f(x).$$

Let $I = \{i \mid \lambda_i^* \neq 0\} \subseteq \{1, 2, \ldots, m\}$, where λ^* is the corresponding Lagrange multiplier. For $k = 1, 2, \ldots$, we introduce the penalty function

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

where α is some positive scalar. We can choose $\epsilon > 0$ small enough such that $f(x^*) \le f(x)$ for all feasible x in the closed sphere $S = \{x \mid ||x - x^*|| \le \epsilon\}$. Let x^k be the unique optimal solution of the problem

$$\min_{x \in S} F^k(x).$$

(An optimal solution exists because of the Weierstrass' theorem, and it is unique because F^k is strongly convex.) In the same way as in the Section 3.1.1, it can be shown that the sequence $\{x^k\}$ converges to x^*, and that $\{kh(x^k)\}$ converges to the vector λ^*. Then, for all $i \in I$ the sequence $\{kh_i(x^k)\}$ converges to $\lambda_i^* \ne 0$. Let N be any neighborhood of x^*. For any $i \in I$, there is an index k_i such that

$$x^k \in N \quad \text{and} \quad \lambda_i^* h_i(x^k) > 0, \qquad \forall\, k \ge k_i.$$

Let $\hat{k} = \max\{k_i \mid i \in I\}$. Then $x^{\hat{k}}$ is in the neighborhood N and $\lambda_i^* h_i(x^{\hat{k}}) > 0$ for all $i \in I$, i.e. $x^{\hat{k}}$ is the desired point.

3.1.12

Let x^* be a regular point and a local minimum for the problem

$$\min_{h(x)=0} f(x),$$

with $f(x^*) \ne 0$. For the constraint set, we have $\{x \mid h(x) = 0\} = \{x \mid ||h(x)|| = 0\} = \{x \mid ||h(x)||^2 = 0\}$. Therefore x^* is also a local minimum for the modified problem

$$\min_{||h(x)||^2=0} f(x).$$

However, none of the feasible points of the original problem is a regular point for the modified problem because $\nabla(||h(x)||^2) = 2\nabla h(x)h(x) = 0$ for all x with $h(x) = 0$. If there existed a Lagrange multiplier λ^* for the local minimum x^* of the modified problem, then we would have

$$\nabla f(x^*) + 2\lambda^* \nabla h(x^*)h(x^*) = \nabla f(x^*) = 0,$$

which contradicts the fact that $\nabla f(x^*) \ne 0$. Hence, the local minimum x^* of the modified problem has no Lagrange multiplier.