Numerical Methods II

Problem Set 7

due in class, May 5, 2004

The following set of questions concerns the quadratic penalty method for solving the equality-constrained minimization problem

$$\text{minimize } f(\boldsymbol{x}) \tag{1a}$$
$$\text{subject to } \boldsymbol{h}(\boldsymbol{x}) = 0 \tag{1b}$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ and $h\colon \mathbb{R}^n \to \mathbb{R}^q$. We assume throughout that f and h are at least twice continuously differentiable.

Recall that $\boldsymbol{x}$ is called a *regular point* if $\boldsymbol{h}(\boldsymbol{x}) = 0$ and the $q \times n$ matrix $\boldsymbol{\nabla h}(\boldsymbol{x})$ has full rank.

Necessary Condition Suppose $\boldsymbol{x}^* \in \mathbb{R}^n$ solves problem (1). Then

(i) there exists a vector of Lagrange multipliers $\boldsymbol{\lambda}^* \in \mathbb{R}^q$ such that

$$\boldsymbol{\nabla} f(\boldsymbol{x}^*) + \boldsymbol{\nabla h}(\boldsymbol{x}^*)^T \boldsymbol{\lambda}^* = 0\,. \tag{2}$$

If $\boldsymbol{x}^*$ is a regular point, then the vector of Lagrange multipliers is unique.

Sufficient Condition Suppose $\boldsymbol{x}^* \in \mathbb{R}^n$ satisfies the constraint $\boldsymbol{h}(\boldsymbol{x}^*) = 0$, condition (i), and

(ii) the Hessian matrix $H = \operatorname{Hess}(f + \boldsymbol{h}^T \boldsymbol{\lambda}^*)(\boldsymbol{x}^*)$ is positive definite on the tangent plane to the constraint manifold. In other words,

$$\boldsymbol{x}^T H \boldsymbol{x} > 0 \qquad \text{for all } \boldsymbol{x} \neq 0 \text{ with } \boldsymbol{\nabla h}\, \boldsymbol{x} = 0\,. \tag{3}$$

Then $\boldsymbol{x}^*$ is a strict local minimizer of problem (1).

Penalty Method The penalty method for solving (1) is the following. Solve the unconstrained minimization problem

$$p_\alpha(\boldsymbol{x}) = f(\boldsymbol{x}) + \frac{\alpha}{2}\, \|\boldsymbol{h}(\boldsymbol{x})\|^2 \tag{4}$$

for a fixed value of α. If the constraint is not satisfied to sufficient accuracy, increase α and repeat.

1. *Monotonicity of the penalty method.* Let $0 < \alpha < \beta$, and let $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ denote the minimizers of p_α and p_β, respectively. Prove the following statements.

   (a) $p_\alpha(\boldsymbol{x}^*) \leq p_\beta(\boldsymbol{y}^*)$

(b) $\|\boldsymbol{h}(\boldsymbol{x}^*)\| \geq \|\boldsymbol{h}(\boldsymbol{y}^*)\|$
Hint: Show that $p_\alpha(\boldsymbol{x}^*) \leq p_\alpha(\boldsymbol{y}^*)$ and $p_\beta(\boldsymbol{y}^*) \leq p_\beta(\boldsymbol{x}^*)$. Add both inequalities and off you go.

(c) $f(\boldsymbol{x}^*) \leq f(\boldsymbol{y}^*)$.

2. *Ill-conditioning of the penalty method.* Show that the minimization problem for p_α gets increasingly ill-conditioned as α becomes large.

   Hint: You have to show that $\operatorname{Hess} p_\alpha(\boldsymbol{x}^*)$ has eigenvalues of very different magnitude. This is the case if you can find two unit vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ so that $\boldsymbol{u}^T \operatorname{Hess} p_\alpha(\boldsymbol{x}^*)\boldsymbol{u}$ and $\boldsymbol{v}^T \operatorname{Hess} p_\alpha(\boldsymbol{x}^*)\boldsymbol{v}$ have very different magnitude.

3. *Continuous dependence of the solution of the penalty parameter.* Let $\mu = 1/\alpha$ and write out the necessary condition for a minimizer of p_α as

$$F(\boldsymbol{x}, \boldsymbol{\lambda}) = \begin{pmatrix} \boldsymbol{\nabla}^T f + \boldsymbol{\nabla}^T \boldsymbol{h}^T \boldsymbol{\lambda} \\ \boldsymbol{h} - \mu \, \boldsymbol{\lambda} \end{pmatrix} = 0 \,.$$

   Show that, provided that the solution $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ satisfies the sufficient condition (ii), and provided that $\boldsymbol{x}^*$ is a regular point, that $\boldsymbol{x}^*$ and $\boldsymbol{\lambda}^*$ vary continuously with μ near $\mu^* = 0$ by verifying that $\boldsymbol{\nabla}_{(\boldsymbol{x},\boldsymbol{\lambda})} F(\boldsymbol{x}^*, \lambda^*; \mu^*)$ is nonsingular. (I.e., the chief condition of the implicit function theorem is satisfied.)