Homework 3

Tom O'Leary-Roseberry

October 19, 2017

3.

Lemma 1. *Let m be the quadratic function defined by*

$$m(p) = g^Tp + \frac{1}{2}p^TBp$$

where B is any symmetric matrix. Then the following statements are true.

(a) m attains a minimum if and only if B is positive semidefinite and g is in the range of B If B is positive semidefinite, then every p satisfying $Bp = -g$ is a global minimizer of m

(b) m has a unique minimizer if and only if B is positive definite.

Proof. (a)

$\Leftarrow$ Since g is in the range of B, there is a p with $Bp = -g$. For all $w \in \mathbb{R}^n$ we have

$$\begin{aligned} m(p+w) &= g^T(p+w) + \frac{1}{2}(p+w)^TB(p+w) \\ &= (g^T + \frac{1}{2}p^TBp) + g^Tw + (Bp)^Tw + \frac{1}{2}w^TBw \\ &= m(p) + \frac{1}{2}w^TBw + (Bp+g)^Tw \\ &\geqslant m(p) \end{aligned} \tag{1}$$

since B is positive semidefinite. Hence p is a minimizer of m

$\Rightarrow$ Let p be a minimizer of m, Since $\nabla m(p) = Bp + g = 0$, we have that g is in the range of B. Also we have that $\nabla^2 m(p) = B$ is positive semidefinite, giving the result.

(b)

$\Leftarrow$ Same as above except now we note that $w^TBw > 0$ whenever $w \neq 0$

$\Rightarrow$ If B is not positive definite, there is a vector $w \neq 0$ such that $Bw = 0$ hence from the above result we have that $m(p+w) = m(p)$ so the minimizer is not unique, giving a contradiction.

□

Theorem 1. *The vector p^* is a global solution of the trust region problem*

$$\min_p m(p) = f + g^Tp + \frac{1}{2}p^TBp \ \textit{s.t.} \ \|p\|_A \leqslant d$$

if and only if p^* is feasible and there is a scalar $\lambda \geqslant 0$ such that the following conditions are satisfied:

$$(B + \lambda A)^*p = -g \tag{2a}$$

$$\lambda(d - \|p^*\|_A) = 0 \tag{2b}$$

$$(B + \lambda A) \ \textit{is positive semidefinite} \tag{2c}$$

Proof. Assume first that there is a $\lambda \geqslant 0$ such that the conditions (2) are met.

The Lemma above imples that p^* is a global minimum of the quadratic function

$$\hat{m}(p) = g^T p + \frac{1}{2}p^T(B + \lambda A)p \quad = \quad m(p) + \frac{\lambda}{2}p^T Ap$$

Since $\hat{m}(p) \geqslant \hat{m}(p^*)$ we have:

$$m(p) \quad \geqslant \quad m(p^*) + \frac{\lambda}{2}((p^*)^T Ap^* - p^T Ap) \tag{3}$$

Because $\lambda(d - \|p^*\|_A) = 0$ and therefore $\lambda(d^2 - \|p^*\|_A^2) = 0$ we have

$$m(p) \geqslant m(p^*) + \frac{\lambda}{2}(d^2 - p^T Ap)$$

Hence from $\lambda \geqslant 0$ we have that $m(p) \geqslant m(p^*)$ for all p with $\|p\|_A \leqslant d$, therefore p^* is a global minimizer.

For the converse assume that p^* is a global minimizer of $m(p)$ and show that there is a $\lambda \geqslant 0$ that satisfies the conditions (2).

In the case that $\|p\|_A < d$, p^* is an unconstrained minimizer of m and so

$$\nabla m(p^*) \quad = \quad Bp^* + g = 0, \qquad \qquad \nabla^2 m(p) = B \text{ is positive semidefinite}$$

and so all the properties hold for $\lambda = 0$

Assume for the remainder of the proof that $\|p^*\|_A = d$, then (2b) is immediately satisfied and p^* also solves the constrained problem

$$\min m(p) \qquad \text{subject to } \|p\|_A = d$$

By applying optimality conditions for constrained optimization to this problem, we find that ther eis a λ such that the Lagrangian function defined by

$$\mathcal{L}(p, \lambda) = m(p) + \frac{\lambda}{2}(p^T Ap - d^2)$$

has a stationary point at p^*. By setting $\nabla_p \mathcal{L}(p^*, \lambda) = 0$, we obtain

$$Bp^* + g + \lambda Ap^* = 0 \qquad \Rightarrow (B - \lambda AI)p^* = -g$$

so that (2a) holds. Since $m(p) \geqslant m(p^*)$ for any p with $p^T Ap = (p^*)^T Ap^* = d^2$, we have for such vectors p that

$$m(p) \quad \geqslant \quad m(p^*) + \frac{\lambda}{2}((p^*)^T Ap^* - p^T Ap)$$

If we substitute the above expression for g, we obtain after some rearranging that

$$\frac{1}{2}(p - p^*)^T(B + \lambda AI)(p - p^*) \quad \geqslant \quad 0$$

Since the set of directions:

$$\left\{ w : w = \pm \frac{p - p^*}{\|p - p^*\|} \text{ for some } p \text{ with } \|p\|_A = d \right\}$$

is dense on the unit sphere, this suffices to prove (2c).

It remains to show that $\lambda \geqslant 0$. Because (2a) and (2c) are satisfied by p^*, we have from the Lemma that p^* minimizes m, so (3) holds. Suppose that there rare only negative directions of λ that satisfy (2a) and (2c). Then we have from (3) that $m(p) \geqslant m(p^*)$ whenever $\|p\|_A \geqslant \|p^*\|_A = d$. Since we already know that p^* minimizes m for $\|p\|_A leqd$ it follows that m is in fact a global unconstrained minimizer of m. From the lemma it follows that $Bp = -g$ and B is positive semidefinite. Therefore the conditions (2a) and (2c) are satisfied by $\lambda = 0$ which contradicts our assumption that only negative values of λ can satisfy the conditions. We conclude that $\lambda \geqslant 0$ completing the proof.

$\square$