SOLUTION MANUAL

7-10-2004

1) The gradient is: $g=\begin{pmatrix}2x+\beta y+1\\2y+\beta x+2\end{pmatrix}$. First order necessary optimality condition $g^*=0$ yields: $\begin{pmatrix}x^*\\y^*\end{pmatrix}=\frac{1}{4-\beta^2}\begin{pmatrix}2\beta-2\\\beta-4\end{pmatrix}$, hence $f(x^*,y^*)=\frac{2\beta-5}{4-\beta^2}$. The Hessian is: $\begin{bmatrix}2&\beta\\\beta&2\end{bmatrix}$ with eigenvalues $\lambda_{1,2}=2\pm\beta$. It is positive definite if $|\beta|<2$, in which case we have a minimum. For $\beta=\pm2$, $g^*=0$ cannot be satisfied and hence no extrema exist. In all other cases the extremum is saddle.

2) First order necessary optimality condition $g^*=0$ yields: $2\left(\sum_k a_k^2\right)x_i^*=2a_i\left(\sum_k a_k x_k^*\right), i=1,2,...,n$. It follows that: $x_i^*=a_i\frac{x_1^*}{a_1}$. Substituting these values we get $f^*=0$. The Hessian is: $\left[a^Ta\ I-aa^T\right]$ and since the second term is of rank one with eigenvalue a^Ta, it follows that the Hessian is positive semidefinite which indicates a minimum.

3) $H'=H-\frac{Hcc^TH}{c^THc}$. (Positive semidefinite). $x=x_0-\lambda H'g$ satisfies the constraint $c^Tx=c^Tx_0=b$, since $c^TH'=0$.

4) $p^Tg<0$ since it is a direction of descent. This means that one can write: $p=-ag+b$, where the vector *b* is normal to *g*, and $a>0$. $(b^Tg=0)$. Substituting we get: $H=I+(a-1)\frac{gg^T}{g^Tg}+\frac{bb^T}{b^Tb}$. Let a vector $s\neq0$. Prove that $s^THs>0$. This is obvious if one notes that $s^Ts-\frac{(s^Tg)^2}{g^Tg}\geq0$.

5)

a. $d=-\nabla f(x)\Rightarrow d=-(Qx+p)$. Since it passes from the origin we have: $d=-p$. The equation of a line passing from the origin along this direction is: $x(\lambda)=-\lambda p,\quad \lambda>0$. Hence the function becomes: $\frac{1}{2}\lambda^2p^TQp-\lambda p^Tp$. Minimizing that wrt λ yields: $\lambda=\frac{p^Tp}{p^TQp}$ hence $x_c=-\frac{p^Tp}{p^TQp}p$, (the Cauchy point).

b. The Newton point is: $x_N=-Q^{-1}p$. The path from the Cauchy to the Newton point is given by: $x(\mu)=x_c+\mu(x_N-x_c),\quad 0\leq\mu\leq1$. Substituting this in the quadratic function and taking the derivative wrt μ we get: $\frac{\partial f}{\partial\mu}=\mu(x_N-x_c)^TQ(x_N-x_c)+(x_N-x_c)^TQx_c+(x_N-x_c)^Tp$ which is an ascending function of μ. Its maximum is at μ*=1*, and its value is vanishing. Hence the gradient wrt μ is negative.