CAAM 454/554 Spring 2013 HW5 solutions

(Contributors: Xin Wang, Ryan George, Jorge Castañón)

March 4, 2013

Problem 2.1 (Nocedal and Wright)

We have $f = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$. To take the gradient, we partially differentiate with respect to each variable:

$$\begin{aligned}
\nabla f(x) &= \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \end{bmatrix} \\
&= \begin{bmatrix} 200(x_2 - x_1^2)(-2x_1) + 2(x_1 - 1) \\ 200(x_2 - x_1^2) \end{bmatrix} \\
&= \begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200(x_2 - x_1^2) \end{bmatrix}
\end{aligned}$$

For the Hessian, we take further partial derivatives :

$$\begin{aligned}
\nabla^2 f(x) &= \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} & \frac{\partial^2}{\partial x_1 \partial x_2} \\ \frac{\partial^2}{\partial x_2 \partial x_1} & \frac{\partial^2}{\partial x_2^2} \end{bmatrix} \\
&= \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}
\end{aligned}$$

We are considering this function in the unconstrained space. Thus by first order necessity, at a minimizer x^*,

$$0 = \nabla f(x^*) = \begin{bmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200(x_2 - x_1^2) \end{bmatrix}$$

From the second row equation $200(x_2 - x_1^2) = 0$, we get that $x_2 = x_1^2$. Thus by the top equation,

$$0 = 400x_1^3 - 400x_1x_2 + 2x_1 - 2 = 400x_1^3 - 400x_1^3 + 2x_1 - 2 = 2x_1 - 2$$

Clearly $[x_1 \;\; x_2]^T = [1 \;\; 1]^T$ is the only point that satisfies first-order necessity. At this point, the Hessian is:

$$\nabla^2 f(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

This has eigenvalues of this matrix are:

```
>> eig([802 -400; -400 200])
ans =
   1.0e+03 *
   0.000399360767488
   1.001600639232512
```

Hence the matrix is positive definite, and the point is a minimum.

Problem 2.2 (Nocedal and Wright)

We calculate the gradient: $\nabla f(x) = \begin{pmatrix} 2x_1 + 8 \\ -4x_2 + 12 \end{pmatrix}$. Setting this to zero gives us the stationary point $[x_1^* \ \ x_2^*]^T = [-4 \ \ 3]$. Since these equations are linear, they admit only one solution, hence there is only one stationary point. The Hessian, evaluated at this point, is:

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$$

Clearly this is not positive definite, for $[0,1]\nabla^2 f(x)[0,1]^T = -4$. Note also that $[10]\nabla^2 f(x)[10]^T = 2$. Thus x^* is a saddle point, not a minimum.

Problem 2.9 (Nocedal and Wright)

The gradient of f at $[1 \ \ 0]^T$ is:

$$\nabla f(x) = \begin{pmatrix} 2(x_1 + x_2^2) \\ 2(x_1 + x_2^2)2x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

Then $\nabla f(x)^T p = (2 \ \ 0)(-1 \ \ 1)^T = -2$. Thus p is a descent direction of f at the point. We now seek the minimum of:

$$\begin{aligned} \left[\min_{\alpha > 0}\right] \phi(\alpha) = f(x + \alpha p) &= f((1,0)^T + \alpha(-1,1)^T) \\ &= f((1-\alpha, \alpha)^T) \\ &= \left(1 - \alpha + \alpha^2\right)^2 \end{aligned}$$

We see that

$$\phi'(\alpha) = 2\left(1 - \alpha + \alpha^2\right)(2\alpha - 1) = 4\alpha^3 - 6\alpha^2 + 6\alpha - 2$$

Matlab's `roots` indicates that $\alpha = .5$ is the only real zero of this function:

```
roots([4 -6 6 -2])
ans =
   0.5000 + 0.8660i
   0.5000 - 0.8660i
   0.5000
```

We thus test this to which second-order conditions are satisfied:

$$\phi''(\alpha) = 12\alpha^2 - 12\alpha + 6$$

$$\phi''(1/2) = 12(1/2)^2 - 12(1/2) + 6 = 12(1/4) = 3 > 0$$

Thus $\alpha = .5$ is the minimum.

Problem 2.13 (Nocedal and Wright)

Show that the sequence $x_k = 1/k$ is not Q-linearly convergent, though it does converge to zero. (This is called sublinear convergence.)

Proof. Obviously, $\lim_{k \to \infty} x_k = 0 = x^*$. However,

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} = \frac{k}{k+1} \to 1, \quad k \to \infty.$$

We can't find a constant $r \in (0, 1)$, such that

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} \leq r, \quad for\ all\ k\ sufficiently\ large.$$

Hence, x_k is not Q-linearly convergent. $\square$

Problem 2.14 (Nocedal and Wright)

Show that the sequence $x_k = 1 + (0.5)^{2^k}$ is Q-quadratically convergent to 1.

Proof. First, it is easy to see $x_k \to 1,\ k \to \infty$. Then

$$\frac{|x_{k+1} - 1|}{|x_k - 1|^2} = \frac{0.5^{2^(k+1)}}{(0.5^{2^k})^2} = 1.$$

By definition, x_k is Q-quadratically convergent to 1. $\square$

Problem 2.15 (Nocedal and Wright)

Does the sequence $1/(k!)$ converge Q-superlinearly? Q-quadratically?

Proof. First, $1/(k!) \to 0,\ k \to \infty$. By the fact that $\frac{1/(k+1)!}{1/k!} = \frac{1}{k+1} \to 0,\ k \to \infty$, the sequence $1/(k!)$ converge Q-superlinearly. However,

$$\frac{1/(k+1)!}{1/(k!)^2} = \frac{k!}{k+1} \to \infty,\ k \to \infty.$$

So, the sequence $1/(k!)$ does not converge Q-quadratically. $\square$