HW2 M529 – Spring 14

Solutions

1. a) Since $g(a)g(b) < 0$, then 0 must be a number between $g(a)$ and $g(b)$:

$$g(a) < 0 < g(b) \quad \text{or}$$

$$g(a) > 0 > g(b)$$

Then, by the IVT there is a number $r \in (a,b)$ such that

$$g(r) = 0$$

To show that r is unique, consider the possibility that there are in fact at least two solutions $r_1, r_2 \in (a,b)$ so that $g(r_1) = 0$ and $g(r_2) = 0$. If this is true, then by Rolle's Theorem there must be a point $c \in (r_1, r_2)$ such that $f'(c) = 0$.

However, this is not possible because

$|g'(x)| \geq m \quad \forall x \in (a,b)$ and therefore $g'(x)$ cannot be zero anywhere in (a,b).

Therefore the solution r is unique.

b) Newton's method formula is

$$x_{k+1} = x_k - \frac{g(x_k)}{g'(x_k)}$$

subtracting r from both sides:

$$x_{k+1} - r = x_k - r - \frac{g(x_k)}{g'(x_k)} \qquad (1)$$

Using Taylor's formula:

$$g(x) = g(x_k) + g'(x_k)(x - x_k) + \frac{1}{2} g''(z)(x - x_k)^2$$

where $z \in [x, x_k]$.

Evaluating at r:

$$g(r) = g(x_k) + g'(x_k)(r - x_k) + \frac{1}{2} g''(z)(r - x_k)^2 = 0$$

therefore

$$g(x_k) = -g'(x_k)(r - x_k) - \frac{1}{2} g''(z)(r - x_k)^2 \qquad (2)$$

Substituting (2) in (1):

$$x_{k+1} - r = x_k - r + \frac{g'(x_k)(r - x_k) + \frac{1}{2} g''(z)(r - x_k)^2}{g'(x_k)}$$

$$= \cancel{x_k - r + (r - x_k)}^{\to 0} + \frac{g''(z)}{2g'(x_k)}(r - x_k)^2$$

$$g'(x_k)(x_{k+1} - r) = \frac{1}{2} g''(z)(x_k - r)^2$$

Taking absolute values:

$$|g'(x_k)|\,|x_{k+1} - r| = \frac{1}{2}|g''(z)|\,|x_k - r|^2$$

Since $|g'(x)| \geq m$:

$$|g'(x_k)|\,|x_{k+1} - r| \geq m\,|x_{k+1} - r|$$

Also, since $|g''(x)| \leq M$, then

$$\frac{1}{2}|g''(z)|\,|x_k - r|^2 \leq \frac{1}{2} M\,|x_k - r|^2$$

$\therefore$

$$m\,|x_{k+1} - r| \leq \frac{M}{2}|x_k - r|^2 \Rightarrow$$

$$|x_{k+1} - r| \leq \frac{M}{2m}|x_k - r|^2$$

2. To show that $p_k = -\nabla f(x_k)$ is the direction of maximum decrease of f, we need to show that

$$p_k^T \nabla f(x_k) \text{ is minimum.}$$

Substituting p_k:

$$-\nabla f(x_k)^T \nabla f(x_k) = \|-\nabla f(x_k)\| \, \|\nabla f(x_k)\| \cos\theta$$

where θ is the angle between $-\nabla f(x_k)$ and $\nabla f(x_k)$.

It is clear that $\theta = \pi$, so $\cos\theta = -1$ (and there is no smaller value for $\cos\theta$), thus:

$$p_k^T \nabla f(x_k) = -\|\nabla f(x_k)\|^2$$

and this value is minimum.

3. If $p_k = -Hf(x_k)^{-1} \nabla f(x_k)$ is a descent direction, then

$$p_k^T \nabla f(x_k) < 0, \text{ or } \nabla f(x_k)^T p_k < 0$$

Thus

$$\nabla f(x_k)^T \left[-Hf(x_k)^{-1} \nabla f(x_k) \right] =$$

$$-\nabla f(x_k)^T Hf(x_k)^{-1} \nabla f(x_k) \quad (1)$$

(1) is going to be negative only if

$$\nabla f(x_k)^T Hf(x_k)^{-1} \nabla f(x_k)$$

is positive. This can be ensured if $Hf(x_k)^{-1}$ is positive definite. Therefore, we can show that p_k is a descend direction if we can show that $Hf(x_k)^{-1}$ is positive definite.

Recall that we can write any symmetric square matrix A as follows:

$$A = P^T D P$$

where P is formed by using the eigenvectors of A and D is a diagonal matrix with the eigenvalues of A in the diagonal, and $P^T = P^{-1}$.

Thus

$$A^{-1} = (P^T D P)^{-1} = (DP)^{-1}(P^T)^{-1} = (DP)^{-1} P =$$
$$P^{-1} D^{-1} P = P^T D^{-1} P$$

Therefore, if A is positive definite, so is A^{-1} because the eigenvalues of this last matrix are simply the reciprocals of the eigenvalues of A

and therefore maintain their sign.

4. Let $B = A + \mu I$, then using $A = P^T D P$ as in the previous exercise:

$$B = P^T D P + \mu I \qquad (1)$$

but since $P^T = P^{-1}$, then $P^T P = I$. Using this result in (1):

$$B = P^T D P + P^T (\mu I) P$$
$$= P^T (D + \mu I) P$$

If D had any negative eigenvalue, the new diagonal matrix

$$D + \mu I$$

will have only positive values because

$$\mu > \max(|\lambda_1|, |\lambda_2|, \cdots, |\lambda_n|)$$

and so if λ_i is the most negative eigenvalue, then the ith entry in the diagonal will have at least a number equal to

$$\overbrace{\lambda_i + |\lambda_i|}^{\geq 0} + \varepsilon > 0$$

5. In Newton's method code, before calculating d (the step's direction):

```
...
l = min(eig(H));
if l < 0
    H = H + (abs(l) + 0.001) * eye(2,2);
end
...
```