STAT 221: STATISTICAL COMPUTING METHODS

Spring, 2004

Solution keys of ASSIGNMENT 1 — Due on Mar. 3, 2004

1.

- Based on the definition of Newton's method, it is not hard to show the given form.
- The right hand side of the iterative form has a minimum at $x_{n-1} = c^{1/m}$ and the minimum is $c^{1/m}$ where $c > 0$. Thus, $x_n \geq c^{1/m}$ for all $x_{n-1} > 0$.
- Since $x_{n-1} \geq c^{1/m}$, we have
$$1 - \frac{1}{m} + \frac{c}{mx_{n-1}^m} \leq 1 - \frac{1}{m} + \frac{c}{mc^{m/m}} = 1,$$
which leads to $x_n \leq x_{n-1}$.
- Letting $g(x_{n-1}) = x_n - x_{n-1}$, we can show that $g'(x_{n-1}) < 0$ whenever $x_{n-1} \geq c^{1/m}$, i.e., x_n monotonely decrease to $c^{1/m}$, which is the minimum value.
- Based on the first result, x_1 will be greater than $c^{1/m}$ when $0 < x_0 < c^{1/m}$, but thereafter, x_n will monotonely decrease to $c^{1/m}$ because of $x_1 > c^{1/m}$.

2.

(a) In order to converge a positive solution of the equation, $|f'(x)| < 1$ should be satisfied. After setting $g(x) = 3x^2 - e^x$ at the first step, we know the first solution of this equation will be between 0 and 1 because of $g(0) < 0$ and $g(1) > 0$, and the second solution between 3 and 4 because of $g(3) > 0$ and $g(4) < 0$. In the vicinity of both roots, however, it is easy to confirm $|f'(x_\infty)|$ is greater than 1.

Thus, we let $f(x) = cg(x)+x$, then try to find a value of c such that $-2 < cg'(x) < 0$ is satisfied. First, since $g'(x)$ is greater than 0 in the vicinity of the first solution, we have $0 < g'(x) < -2/c$ where c is negative, which means c should be between $-1/(3\log 6 - 3)$ and 0 because $g'(x)$ has an upper bound of $(6\log 6 - 6)$. As long as c is set to a number somewhere in this interval, we guarantee that fixed-point iteration will converge to a positive (first) solution on an interval that satisfies $g'(x) > 0$. Finding the interval gives birth to another non-linear equation problem, but this can be done easily by using the bisection method. For instance, when we set $c = -1/4$, the solution to this equation is $x^* = 0.9100076$ and the corresponding interval on which the fixed-point iteration converges is (0.2044814, 2.833148).

Second, $g'(x)$ is less than 0 in the vicinity of the second solution, and thus we have $-2/c < g'(x) < 0$ where c is positive. Because $g'(x)$ does not have a lower bound, the interval that converges to a positive (second) solution will hinge on the value of c. When we set $c = 1/100$, the solution to this equation is $x^* = 3.733079$ and the corresponding interval on which the fixed-point iteration converges is (2.833148, 5.449743).

(b) For our first guess, we let $f(x) = 2x - \cos x$ after setting $g(x) = x - \cos x$. However, $|f'(x)| < 1$ cannot be satisfied for any value of x. Thus, we again let $f(x) = cg(x) + x$, then try to find c that satisfies the proposition. It is easy to see when $c \in (-1, 0)$, $|f'(x)| < 1$ is always satisfied for all x except odd multipliers of π plus $\pi/2$. The solution to this equation is $x^* = 0.7390851$.

3.

The likelihood function can be viewed as

$$\prod_{i=1}^{n} \left\{ \left(pf(t_i)\right)^{z_i} \left(1 - pF(t_i)\right)^{1-z_i} \right\},$$

then the loglikelihood is given by

$$\ell = \sum_{i=1}^{n} \left\{ z_i \log\left(pf(t_i)\right) + (1 - z_i) \log\left(1 - pF(t_i)\right) \right\}.$$

Thus, the first and second derivatives are given by

$$\begin{aligned} \frac{\partial \ell}{\partial p} &= \frac{\sum_{i=1}^{n} z_i}{p} - \sum_{i=1}^{n} \frac{(1 - z_i)F(t_i)}{1 - pF(t_i)} \\ \frac{\partial^2 \ell}{\partial p^2} &= -\frac{\sum_{i=1}^{n} z_i}{p^2} - \sum_{i=1}^{n} \frac{(1 - z_i)F(t_i)^2}{\left(1 - pF(t_i)\right)^2}. \end{aligned}$$

Then we find a solution that satisfies $\ell'(p) = 0$ by using Newton-Raphson and bisection as follows:

```
risk <- read.table("c:/Splus/Data/risk.dat",header=T)
z <- (risk[,2]-40)/15 # We standardize the data.

ell.f <- function(p){ sum(risk[,1])/p-sum((1-risk[,1])*pnorm(z)/(1-p*pnorm(z))) }
ell.s <- function(p){ -sum(risk[,1])/p^2-sum((1-risk[,1])*pnorm(z)^2/(1-p*pnorm(z))^2) }

### Bisection ###
a <- .4
b <- .6
ell.f(a)*ell.f(b) # Is this less than 0?
while(abs(a-b)>1.0e-8){
   c <- (a+b)/2
   if(ell.f(c)==0){
      print("optimal value!")
      break
   }
   else{
      if(ell.f(a)*ell.f(c)<0) b <- c
      else a <- c
   }
}
c # 0.5058496

### Newton-Raphson ###
x <- .4
while(abs(ell.f(x))>1.0e-8) x <- x - ell.f(x)/ell.s(x)
x # 0.5058496

sqrt(-1/ell.s(x)) # 0.0704326 which is the standard error of p.hat.
```

4.
For any non-negative integers n and m where $n > m$,

$$\|B_n - B_m\| \;=\; \left\| \sum_{k=m+1}^{n} \frac{A^k}{k!} \right\| \;\leq\; \sum_{k=m+1}^{n} \frac{\|A^k\|}{k!} \;\leq\; \sum_{k=m+1}^{n} \frac{\|A\|^k}{k!} \;<\; \sum_{k=m+1}^{\infty} \frac{\|A\|^k}{k!}$$

and the right-hand side will converge to 0 as $m \to \infty$. Thus B_n is a Cauchy sequence, so it will converge in the reals.

In particular, if we assume A is a $p \times p$ symmetric matrix, then by an eigen decomposition A can be decomposed into QDQ^{-1} where D is a diagonal matrix whose entries are eigenvalues and Q is a matrix of corresponding eigenvectors. Thus, the given series can be rewritten as

$$B_n \;=\; \sum_{k=0}^{n} \frac{A^k}{k!} \;=\; Q\left(\sum_{k=0}^{n} \frac{D^k}{k!}\right)Q^{-1} \;=\; Q \begin{vmatrix} \sum_k \frac{d_1^k}{k!} & 0 & \cdots & 0 \\ 0 & \sum_k \frac{d_2^k}{k!} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_k \frac{d_p^k}{k!} \end{vmatrix} Q^{-1},$$

and as $n \to \infty$ this will converge to

$$Q \begin{vmatrix} e^{d_1} & 0 & \cdots & 0 \\ 0 & e^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{d_p} \end{vmatrix} Q^{-1} \;\equiv\; Qe^{D}Q^{-1}.$$

5.

- $1 = \|I\| = \|AA^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = \mathrm{cond}(A)$.
- $\mathrm{cond}(A^{-1}) = \|A^{-1}\| \cdot \|A\| = \|A\| \cdot \|A^{-1}\| = \mathrm{cond}(A)$.
- $\mathrm{cond}(cA) = |c| \cdot \|A\| \cdot |c^{-1}| \cdot \|A^{-1}\| = \|A\| \cdot \|A^{-1}\| = \mathrm{cond}(A)$.
- $\mathrm{cond}_2(U) = \|U\|_2 \cdot \|U^{-1}\|_2 = \sqrt{\rho(U^\top U)}\sqrt{\rho(U^{-\top}U^{-1})} = \sqrt{\rho(U^\top U)}\sqrt{\rho(UU^\top)} = 1$.
- $\mathrm{cond}_2(AU) = \|AU\|_2 \cdot \|U^{-1}A^{-1}\|_2 = \|U^\top A^\top\|_2 \cdot \|U^{-1}A^{-1}\|_2 = \|A^\top\|_2 \cdot \|A^{-1}\|_2 = \mathrm{cond}_2(A)$.
- $\mathrm{cond}_2(UA) = \|UA\|_2 \cdot \|A^{-1}U^{-1}\|_2 = \|UA\|_2 \cdot \|U^{-\top}A^{-\top}\|_2 = \|A\|_2 \cdot \|A^{-\top}\|_2 = \mathrm{cond}_2(A)$.