

Problem Set 7

github: stephaniesookim

November 14, 2014

/raggedright

Number 1

I apologize for late submittance of number 1. I was unaware that I should have submitted by monday.

1. What are the goals of their simulation study and what are the metrics that they consider in assessing their method? The purpose of the simulation study is first to assess the accuracy of the proposed asymptotic approximation in finite samples and second to examine the power of the EM test.

The metrics they considered are the power of the EM and type I error.

2. What choices did the authors have to make in designing their simulation study? What are the key aspects of the data generating mechanism that likely affect the statistical power of the test? The authors made choices of the sample sizes (200 and 400), the significance levels (5% and 1%), the number of repetitions (1000), the number of iterations, etc.

The key aspects of the data generating mechanism that are likely to affect the statistical power of the test are the sample size, the component means, variances.

3. Interpret their tables on power (Tables 4 and 6) - do the results make sense in terms of how the power varies as a function of the data generating mechanism? From both tables 4 and 6, we can observe that the power is greater when the sample size is larger and m is larger. The results make sense.

Number 2

Submitting by hard-copy

Number 3 (a)

The most efficient way is to multiply the diagonal entries of D and X

```
«eval=FALSE»
```

```
X <- matrix(1:30, 5, 6) D <- diag(1:5) DX <- diag(D) * X DX
```

```
@
```

Number 3 (b)

The most efficient way is to use transpose of X and diagonal entries and take a transpose of it again

```
«eval=FALSE»
```

```
X <- matrix(1:30, 5, 6) D <- diag(1:5) XD <- t(t(X) * diag(D)) XD
```

```
@
```

Number 4

(a)

The number of flops involved in the forward reduction step of the LU decomposition can be counted step by step.

The first step produces 0's of which each 0 requires 1 multiplication and 1 subtraction.

Thus there will be n^2 flops.

Similarly, in the second step there will be $(n-1)^2$ flops.

Thus, the total number of flops can be calculated as $n^2 + (n-1)^2 + \dots + 2^2 + 1^2$ which is approximately $1/3 * n^3$.

Cholesky decomposition requires $O(n^3) + (1/6 * n^3)$ calculations.

Thus Cholesky is more efficient

(b)

$1+2+\dots+(n-1)+n = 1/2 * n(n+1)$ flops

(c)

The flops will be added by p in (a) while it will be multiplied by p in (b)

Number 5

We calculate beta hat by using qr, backsolve, crossprod functions.

```
beta_hat <- function(X, Y, A, b) {  
  X.qr = qr(X)  
  R = qr.R(X.qr)  
  Q = qr.Q(X.qr)  
  first = backsolve(R, crossprod(Q, Y))  
  second = tcrossprod(solve(crossprod(R)), A)  
  third = solve(A %*% solve(R) %*% t(AR_1))  
  fourth = -A %*% first + b  
  betahat = first + second %*% third %*% fourth  
  return(betahat)  
}
```