# Lec 4: Sweep Operator

Ailin Zhang

2022-09-23

# Agenda

- Gauss Jordan wrap up
- The Sweep Operator

# Matrix Decomposition

For linear system, $Ax = b$, one never actually computes $A^{-1}b$.

When $A = \mathbf{X}^\top X$ is symmetric, some special computational methods are available:

- Sweep operator
- QR decomposition
- Cholesky decomposition

# Gauss Jordan

$$\mathrm{GJ}[1:n][A|b] = [I|A^{-1}b] = A^{-1}[A|b],$$
$$\mathrm{GJ}[1:n][A|I] = [I|A^{-1}] = A^{-1}[A|I].$$

# Gauss Jordan

$$\mathrm{GJ}[1:n][A|b] = [I|A^{-1}b] = A^{-1}[A|b],$$
$$\mathrm{GJ}[1:n][A|I] = [I|A^{-1}] = A^{-1}[A|I].$$

Matrix version:

$$\mathrm{GJ}[1:m]\begin{bmatrix} A_{11} & A_{12} & | & I_1 & 0 \\ A_{21} & A_{22} & | & 0 & I_2 \end{bmatrix}$$
$$= \begin{bmatrix} I_1 & A_{11}^{-1}A_{12} & | & A_{11}^{-1} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} & | & -A_{21}A_{11}^{-1} & I_2 \end{bmatrix}$$

# Sweep operator

The sweep operator arises from a close analysis of elimination applied to the equation $\mathbf{X}^\top X\beta = \mathbf{X}^\top \mathbf{Y}$. For a symmetric matrix$A$, $B = Sweep(A, k)$ is obtained as follows:

1. Divide $k^{th}$ row and $k^{th}$ column by $a_{kk}$:

$$b_{ik} = a_{ik}/a_{kk};\ b_{kj} = a_{kj}/a_{kk}$$

2. Subtract $a_{ik}a_{kj}/a_{kk}$ from the other entries:

$$b_{ij} = a_{ij} - a_{ik}a_{kj}/a_{kk}$$

3. Invert $k^{th}$ diagonal element:

$$b_{kk} = -1/a_{kk}$$

# Sweep operator

The sweep operator arises from a close analysis of elimination applied to the equation $\mathbf{X}^\top X\beta = \mathbf{X}^\top\mathbf{Y}$. For a symmetric matrix$A$, $B = Sweep(A, k)$ is obtained as follows:

1. Divide $k^{th}$ row and $k^{th}$ column by $a_{kk}$:

$$b_{ik} = a_{ik}/a_{kk};\ b_{kj} = a_{kj}/a_{kk}$$

2. Subtract $a_{ik}a_{kj}/a_{kk}$ from the other entries:

$$b_{ij} = a_{ij} - a_{ik}a_{kj}/a_{kk}$$

3. Invert $k^{th}$ diagonal element:

$$b_{kk} = -1/a_{kk}$$

Question: What is the time complexity?

# Sweep operator

The sweep operator arises from a close analysis of elimination applied to the equation $\mathbf{X}^\top X\beta = \mathbf{X}^\top \mathbf{Y}$. For a symmetric matrix$A$, $B = Sweep(A, k)$ is obtained as follows:

1. Divide $k^{th}$ row and $k^{th}$ column by $a_{kk}$:

$$b_{ik} = a_{ik}/a_{kk};\ b_{kj} = a_{kj}/a_{kk}$$

2. Subtract $a_{ik}a_{kj}/a_{kk}$ from the other entries:

$$b_{ij} = a_{ij} - a_{ik}a_{kj}/a_{kk}$$

3. Invert $k^{th}$ diagonal element:

$$b_{kk} = -1/a_{kk}$$

Question: What is the time complexity? A single sweep operation visits every element of the matrix and requires roughly one divide per element. This leads to roughly $\mathcal{O}(n^2)$ divides.

# R code

```
mySweep <- function(A, m)
{
n <- dim(A)[1]
for (k in 1:m)
{
for (i in 1:n)
 for (j in 1:n)
    if (i!=k  & j!=k)
       A[i,j] <- A[i,j] - A[i,k]*A[k,j]/A[k,k]
for (i in 1:n)
  if (i!=k)
    A[i,k] <- A[i,k]/A[k,k]
for (j in 1:n)
  if (j!=k)
    A[k,j] <- A[k,j]/A[k,k]
A[k,k] <- - 1/A[k,k]
}
return(A)
}
A = matrix(c(1,2,3,7,11,13,17,21,23), 3,3)
solve(A)
```

```
##       [,1] [,2]  [,3]
## [1,]  1.00 -3.0  2.00
## [2,] -0.85  1.4 -0.65
## [3,]  0.35 -0.4  0.15
```

```
mySweep(A,3)
```

# Proposition 1

### Proposition

*Suppose* $V_{p\times m} = U_{p\times m}A_{m\times m}$, *and* $B = Sweep(A, k)$. *Then* $\hat{V} = \hat{U}B$, *where:*

- $\hat{U} = U$ *except that its* $k^{th}$ *column is* $v_k$;
- $\hat{V} = V$ *except that its* $k^{th}$ *coumn is* $-u_k$.

Proof:

# Proposition 2

**Proposition**

*If $A$ is a symmetric invertible matrix and we sweep on each diagonal element of $A$ the result is* $B = Sweep(A, 1:n) = -A^{-1}$

Proof:

# Proposition 3

**Proposition**

*If* $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ *and we sweep on each of the diagonal entries of* $A_{11}$, *we get* $B = \begin{bmatrix} -A_{11}^{-1} & A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$

Proof: Start with an simple example $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

# Proposition 3

**Proposition**

*If* $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ *and we sweep on each of the diagonal entries of* $A_{11}$, *we get* $B = \begin{bmatrix} -A_{11}^{-1} & A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}$

Proof: Take advantage of proposition 2.

# Proposition 4

We construct a matrix $\mathbf{Z} = [\mathbf{XY}]$, and let

$$A = \mathbf{Z}^\top \mathbf{Z} = \begin{bmatrix} \mathbf{X}^\top \mathbf{X} & \mathbf{X}^\top \mathbf{Y} \\ \mathbf{Y}^\top \mathbf{X} & \mathbf{Y}^\top \mathbf{Y} \end{bmatrix}$$

be the cross-product matrix. Then

$$\mathrm{SWP}[1:p]A = \begin{bmatrix} -(\mathbf{X}^\top \mathbf{X})^{-1} & (\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{X}^\top \mathbf{Y} \\ \mathbf{Y}^\top \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} & \mathbf{Y}^\top \mathbf{Y} - \mathbf{Y}^\top \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{X}^\top \mathbf{Y} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{\mathrm{Var}(\hat{\beta})}{\sigma^2} & \hat{\beta} \\ \hat{\beta}^\top & \mathrm{RSS} \end{bmatrix}$$

where $\mathrm{RSS} = \|\mathbf{Y} - \mathbf{X}\hat{\beta}\|_{\ell_2}^2$ is the residual sum of squares.

# Other Application of sweep operator

Perform sweep operator on multivariate normal distribution

$$\begin{bmatrix} \Omega & x-\mu \\ x^t-\mu^t & 0 \end{bmatrix}$$

# Why Sweep Operator?

1. Efficient for computing the central statistics used in multiple regression
   - Start from a square correlation or covariance matrix.
   - Compute multiple correlation, residual variance, regression slopes, and standard errors of slopes, plus some other
2. An efficient way to compute a whole series of regressions, as in stepwise regression.
3. (Drawback) If it is applied many times to the same matrix, rounding error can accumulate.