# GENE638 - Homework 2

**Daniel Osorio - dcosorioh@tamu.edu**
**Department of Veterinary Integrative Biosciences**
**Texas A&M University**

1. For this system of equations:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 20 \\ 25 \end{bmatrix}$$

   (a) Determine the rank of the coefficient matrix

```
> as.numeric(Matrix::rankMatrix(coefficientMatrix))
[1] 2
```

   (b) Express any linearly dependent columns in the matrix form $\begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A'_{12} \end{bmatrix} L$

$$\begin{aligned}
A &= \left[\begin{array}{c:cc} 1 & 0 & 1 \\ \hdashline 0 & 2 & 2 \\ 1 & 2 & 3 \end{array}\right] \\
A_{\begin{bmatrix} 11 \\ 12 \end{bmatrix}} &= \begin{bmatrix} 0 & 1 \\ 2 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
A_{\begin{bmatrix} 12 \\ 12 \end{bmatrix}} &= \begin{bmatrix} (0 \times -1 + 1 \times 1) \\ (2 \times -1 + 2 \times 1) \\ (2 \times -1 + 3 \times 1) \end{bmatrix} \\
A_{\begin{bmatrix} 12 \\ 12 \end{bmatrix}} &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}
\end{aligned} \tag{1}$$

   (c) Find a generalized inverse of the coefficient matrix. Prove that it is a generalized inverse

```
> gInverse <- MASS::ginv(coefficientMatrix)
> gInverse
           [,1]          [,2]         [,3]
[1,]  0.5000000 -3.333333e-01 1.666667e-01
[2,] -0.3333333  3.333333e-01 1.942890e-16
[3,]  0.1666667  6.938894e-17 1.666667e-01

> round(gInverse %*% coefficientMatrix)
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
```

```
> round(coefficientMatrix %*% gInverse %*% coefficientMatrix)
     [,1] [,2] [,3]
[1,]    1    0    1
[2,]    0    2    2
[3,]    1    2    3
```

(d) Using the inverse from part (c), solve for $x$. Prove your solution satisfies the equations

```
> x <- gInverse %*% c(5, 20, 25)
> round(x)
     [,1]
[1,]    0
[2,]    5
[3,]    5
> coefficientMatrix %*% x
     [,1]
[1,]    5
[2,]   20
[3,]   25
```

(e) What do your solutions estimate?

(f) Based in what you did in part (e): Can you estimate $x_1$? Can you estimate $x_1 - x_2$?

2. Using the partitioned matrix inverse procedure on page 14 in the notes

(a) Find the inverse of: $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

(b) Prove that your answer in (a) is an inverse

(c) Solve $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

(d) What is the solution in (c) an estimate of?

(e) Can you estimate $x_3$? With what? Can you estimate $6x_1 + 4x_2 - 8x_3$? With what?

3. Assuming the linear model $y_i = \mu + \epsilon_i$ where $y_i$ is and observed first lactation milkfat production, $\mu$ is the population mean milkfat production, and $\epsilon_i$ is the deviation of an individual cow's production from the mean. Write this model as

$$\underline{y} = 1\mu + \underline{\epsilon}$$

| Cow | $1^{st}$ lactation milkfat (lb) |
|---|---|
| 1 | 300 |
| 2 | 290 |
| 3 | 405 |
| 4 | 360 |
| 5 | 315 |

(a) Calculate:

i. $X'X$

```
> X <- rep(1,length(y))
> t(X) %*% X
```

```
     [,1]
[1,]    5
```

ii. $(X'X)^{-1}$

```
> solve(t(X) %*% X)
     [,1]
[1,]  0.2
```

iii. $X'\underline{y}$

```
> t(X) %*% y
     [,1]
[1,] 1670
```

(b) Solve $\hat{\mu} = (X'X)^{-1}X'\underline{y}$

```
> muHat <- as.numeric(solve(t(X) %*% X) %*% X %*% y)
> muHat
[1] 334
```

(c) Show that $\hat{\mu} = \overline{y}$

```
> meanY <- mean(y)
> meanY
[1] 334
> all.equal(muHat, meanY)
[1] TRUE
```

(d) Find the predicted deviations $\hat{\underline{\epsilon}} = \underline{y} - \underline{1}\hat{\mu}$

```
> epsilonHat <- (y - muHat)
> epsilonHat
[1] -34 -44  71  26 -19
```

(e) Find $\frac{\hat{\underline{\epsilon}}'\hat{\underline{\epsilon}}}{n-rank(X)}$ and show that this is $s^2$ the sample variance.

```
> s2Hat <- as.numeric((t(epsilonHat) %*% epsilonHat) / (length(y) - 1))
> s2Hat
[1] 2292.5
> var(y)
[1] 2292.5
> all.equal(s2Hat, var(y))
[1] TRUE
```