3 Lecture 3:Jan 25

Last time

- Git
- Linear algebra: vector and vector space, rank of a matrix

Today

- Column space and Nullspace (JM Appendix A)
- Simple Linear Regression JF Chapter 5

Column space

Definition: The column space of a matrix, denoted by $C(\mathbf{A})$ is the vector space spanned by the columns of the matrix, that is,

$$C(\mathbf{A}) = \{\mathbf{x} : \text{ there exists a vector } \mathbf{c} \text{ such that } \mathbf{x} = \mathbf{Ac}\}.$$

This means that if $\mathbf{x} \in C(\mathbf{A})$, we can find coefficients c_j such that

$$\mathbf{x} = \sum_j c_j \mathbf{a}^{(j)}$$

where $\mathbf{a}^{(j)} = \mathbf{A}_{\cdot j}$ denotes the jth column of matrix $\mathbf{A}$.

- The column space of a matrix consists of all vectors formed by multiplying that matrix by any vector.
- The number of basis vectors for $C(\mathbf{A})$ is then the number of linearly independent columns of the matrix $\mathbf{A}$, and so, $\dim(C(\mathbf{A})) = \text{rank}(\mathbf{A})$.
- The dimension of a space is the number of vectors in its basis.

Example A.2

Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & -3 \\ 1 & 2 & -1 \\ 1 & 3 & 1 \\ 1 & 4 & 3 \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$. Show that $\mathbf{Ac}$ is a linear combination of columns in $\mathbf{A}$.

solution:

$$\mathbf{Ac} = \begin{bmatrix} 1 \times 5 + 1 \times 4 + (-3) \times 3 \\ 1 \times 5 + 2 \times 4 + (-1) \times 3 \\ 1 \times 5 + 3 \times 4 + 1 \times 3 \\ 1 \times 5 + 4 \times 4 + 3 \times 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \\ 20 \\ 30 \end{bmatrix}.$$

You could recognize that

$$\mathbf{A}\mathbf{c} = 5 \times \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 4 \times \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + 3 \times \begin{bmatrix} -3 \\ -1 \\ 1 \\ 3 \end{bmatrix} = 5\mathbf{a}^{(1)} + 4\mathbf{a}^{(2)} + 3\mathbf{a}^{(3)} = \begin{bmatrix} 0 \\ 10 \\ 20 \\ 30 \end{bmatrix}.$$

Result A.1

$\text{rank}(\mathbf{AB}) \leqslant \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$.

proof: Each column of $\mathbf{AB}$ is a linear combination of columns of $\mathbf{A}$ (i.e. $(\mathbf{AB})_{\cdot j} = \mathbf{A}\mathbf{b}^{(j)}$), so the number of linearly independent columns of $\mathbf{AB}$ cannot be greater than that of A. Similarly, $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B}^T\mathbf{A}^T)$, the same argument gives $\text{rank}(\mathbf{B}^T)$ as an upper bound.

Result A.2

- (a) If $\mathbf{A} = \mathbf{BC}$, then $C(\mathbf{A}) \subseteq C(\mathbf{B})$.
- (b) If $C(\mathbf{A}) \subseteq C(\mathbf{B})$, then there exists a matrix $\mathbf{C}$ such that $\mathbf{A} = \mathbf{BC}$.

proof: For (a), any vector $\mathbf{x} \in C(\mathbf{A})$ can be written as $\mathbf{x} = \mathbf{A}\mathbf{d} = \mathbf{B}(\mathbf{C}\mathbf{d})$. For (b), $\mathbf{A}_{\cdot j} \in C(B)$, so that there exists a vector $\mathbf{c}^{(j)}$ such that $\mathbf{A}_{\cdot j} = \mathbf{B}\mathbf{c}^{(j)}$. The matrix $\mathbf{C} = (\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \ldots, \mathbf{c}^{(n)})$ satisfies that $\mathbf{A} = \mathbf{BC}$.

Null space

Definition: The null space of a matrix, denoted by $N(\mathbf{A})$, is $N(\mathbf{A}) = \{\mathbf{y} : \mathbf{A}\mathbf{y} = \mathbf{0}\}$.

Result A.3

If $\mathbf{A}$ has full-column rank, then $N(\mathbf{A}) = \{\mathbf{0}\}$.

proof: Matrix $\mathbf{A}$ has full-column rank means its columns are linearly independent, which means that $\mathbf{A}\mathbf{c} = \mathbf{0}$ implies $\mathbf{c} = \mathbf{0}$.

Theorem A.1

Assume $\mathbf{A} \in \mathbb{R}^{m \times n}$, then $\dim(C(\mathbf{A})) = r$ and $\dim(N(\mathbf{A})) = n - r$, where $r = \text{rank}(\mathbf{A})$.

See JM Appendix Theorem A.1 for the proof.

Interpretation: "dimension of column space + dimension of null space = # columns"

*Mis*Interpretation: Columns space and null space are orthogonal complement to each other. They are of different orders in general! Next result gives the correct statement.

Simple linear regression

Figure 3.1 shows Davis's data on the measured and reported weight in kilograms of 101 women who were engaged in regular exercise.

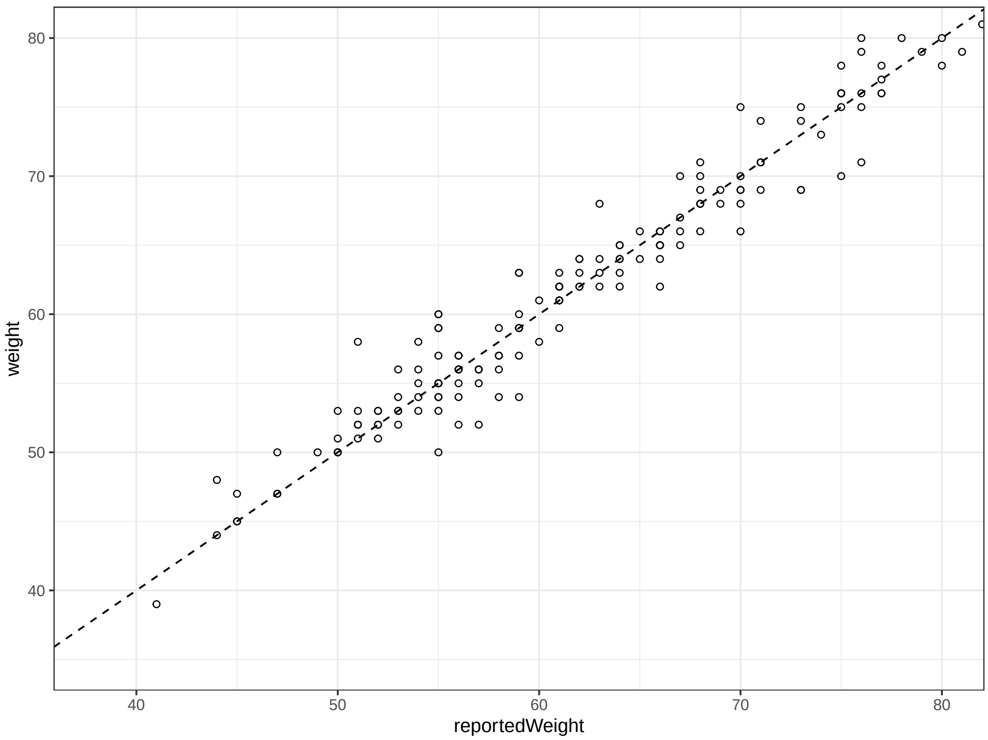

Figure 3.1: Scatterplot of Davis's data on the measured and reported weight of 101 women. The dashed line gives $y = x$.

It's reasonable to assume that the relationship between measured and reported weight appears to be linear. Denote:

- measured weight by y_i: **response variable** or **dependent variable**
- reported weight by x_i: **predictor variable** or **independent variable**
- intercept: β_0
- slope: β_1
- residual/error term ϵ_i.

Then the simple linear regression model writes:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i.$$

For given $(\hat{\beta}_0, \hat{\beta}_1)$ values, the *fitted value* or *predicted value* for observation i is:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i.$$

Therefore, the residual is

$$\epsilon_i = y_i - \hat{y}_i$$

Fitting a linear model

Choose the "best" values for β_0, β_1 such that

$$SS[E] = \sum_1^n \Big(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\Big)^2 = \sum_1^n (y_i - \hat{y}_i)^2 = \sum_1^n \epsilon_i^2$$

is minimized. These are **least squares** (LS) estimates:

$$\begin{aligned} \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \bar{x} \\ \hat{\beta}_1 &= \tfrac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}. \end{aligned}$$

Definition: The line satisfying the equation

$$y = \hat{\beta}_0 + \hat{\beta}_1 x$$

is called the <u>linear regression</u> of y on x which is also called the <u>least squares line</u>.

For Davis's data, we have

$$\begin{aligned} n &= 101 \\ \bar{y} &= \frac{5780}{101} = 57.228 \\ \bar{x} &= \frac{5731}{101} = 56.743 \\ \sum(x_i - \bar{x})(y_i - \bar{y}) &= 4435.9 \\ \sum(x_i - \bar{x})^2 &= 4539.3, \end{aligned}$$

so that

$$\begin{aligned} \hat{\beta}_1 &= \frac{4435.9}{4539.3} = 0.97722 \\ \hat{\beta}_0 &= 57.228 - 0.97722 \times 56.743 = 1.7776 \end{aligned}$$