Matrix Structure

Computational Mathematics and Statistics Camp
University of Chicago
September 2018

1. Recall the general rule for calculating the determinant of an $n \times n$ matrix:

$$|\mathbf{X}| = \sum_{j=1}^{n} (-1)^{i+j} x_{ij} |\mathbf{X}_{[ij]}|$$

where the ijth **minor** of $\mathbf{X}$ for x_{ij}, $|\mathbf{X}_{[ij]}|$, is the determinant of the $(n-1) \times (n-1)$ submatrix that results from taking the ith row and jth column out. The **cofactor** of $\mathbf{X}$ is the minor signed as $(-1)^{i+j} x_{ij} |\mathbf{X}_{[ij]}|$. To calculate the determinant we cycle recursively through the columns and take sums with a formula that multiplies the cofactor by the determining value.

Given this rule, obtain the trace and determinant of the following matrix. You can do this the hard way, or the easy(ier) way. I encourage you to think a bit before starting the calculations for the determinant - how can you make the problem easier?

$$\begin{bmatrix} 6 & 6 & 1 & 0 \\ 0 & 4 & 0 & 1 \\ 4 & 2 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

2. Invert each of the following matricies. Verify you have the correct inverse by calculating $\mathbf{X}\mathbf{X}^{-1} = \mathbf{I}$. Not all of the matricies may be invertible - if not, show why.

   a. $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

   b. $\begin{bmatrix} 4 & 5 \\ 2 & 4 \end{bmatrix}$

   c. $\begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}$

   d. $\begin{bmatrix} 2 & 4 & 0 \\ 4 & 6 & 3 \\ -6 & -10 & 0 \end{bmatrix}$

3. When it comes to real numbers, we know that if $xy = 0$, then either $x = 0$ or $y = 0$ or both. One might believe that a similar idea applies to matricies, but one would be wrong. Prove that if the matrix product $\mathbf{AB} = \mathbf{0}$ (by which we mean a matrix of appropriate dimensionality made up entirely of zeroes), then it is not necessarily true that either $\mathbf{A} = \mathbf{0}$ or $\mathbf{B} = \mathbf{0}$. Hint: in order to prove that something is not always true, simply identify one example where $\mathbf{AB} = \mathbf{0}$, $\mathbf{A}, \mathbf{B} \neq \mathbf{0}$.

4. Ordinary least squares regression is a common method for obtaining regression parameters relating a set of explanatory variables with a continuous outcome of interest. The vector $\hat{\mathbf{b}}$ that contains the intercept and the regression slope is calculated by the equation:

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

If an explanatory variable is nominal (i.e. ordering does not matter) with more than two classes (e.g. $\{\text{White}, \text{Black}, \text{Asian}, \text{Mixed}, \text{Other}\}$), the variable must be modified to include in the regression model. A common technique known as **one-hot encoding** converts the column into a series of $n-1$ binary

(0/1) columns where each column represents a single class and n is the total number of unique classes in the original column. Explain why this method converts the column into $n-1$ columns, rather than n columns, in terms of linear algebra.

5. Solve the following systems of equations for x, y, z, either via matrix inversion or substitution:

   a. System #1
   $$\begin{aligned} x + y + 2z &= 2 \\ 3x - 2y + z &= 1 \\ y - z &= 3 \end{aligned}$$

   b. System #2
   $$\begin{aligned} 2x + 3y - z &= -8 \\ x + 2y - z &= 2 \\ -x - 4y + z &= -6 \end{aligned}$$

   c. System #3
   $$\begin{aligned} x - y + 2z &= 2 \\ 4x + y - 2z &= 10 \\ x + 3y + z &= 0 \end{aligned}$$

6. Recall from Gill 4.8 that every $p \times p$ matrix $\mathbf{X}$ has p scalar values, $\lambda_i, i = 1, \ldots, p$ such that:

   $$\mathbf{X}\mathbf{h}_i = \lambda_i \mathbf{h}_i$$

   for some corresponding vector $\mathbf{h}_i$. λ_i is called an **eigenvalue** of $\mathbf{X}$ and h_i is called an **eigenvector** of $\mathbf{X}$.

   Now consider the matrix:

   $$\mathbf{M} = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & -3 \end{bmatrix}$$

   along with the eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 2$, and eigenvectors:

   $$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

   a. For each eigenvalue/vector pair i, show that $\mathbf{M}\mathbf{x}_i = \lambda_i \mathbf{x}_i$.

   b. One way to calculate the eigenvalues of $\mathbf{A}$ is to find the values of λ that solve the equation:

   $$|\mathbf{A} - \lambda_i \mathbf{I}| = 0$$

   where $\mathbf{I}$ is the identity matrix. Show that this fact holds for $\mathbf{M}$ given above.