Statistics 206: Handouts – Basics about Matrix Algebra and Random Vectors

September 2019

The matrix approach is necessary for multiple regression. Here is a list of definitions and properties of matrix algebra that we need in this course. There is also a list of definitions and properties of the mean vector and variance-covariance matrix of a random vector. At the end of the notes, there are also some exercises (and solutions).

1 Vectors and matrices

- A **matrix** is a set of elements arranged in a rectangular array.
- Example:

$$\begin{bmatrix} 16 & 23 \\ 33 & 47 \\ 21 & 35 \end{bmatrix}$$

  - The **dimension** of this matrix is 3×2, i.e., 3 rows and 2 columns.
  - The $(1,1)$ element of this matrix is 16 and the $(1,2)$ element is 23: the first index identifies the row number and the second index identifies the column number. Question: what is the $(2,2)$ element?
  - The $(1,1)$, $(2,2)$, etc. elements are called **diagonal elements**. Other elements are called **off diagonal elements**.

- **Denote a matrix by symbols**. For example:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

  is a 2×3 matrix with the (i,j)th element being a_{ij}. A more compact notation for matrix $\mathbf{A}$ is

$$\mathbf{A} = [a_{ij}], \quad i = 1, 2; j = 1, 2, 3.$$

- A matrix with r rows and c columns may be denoted by

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2c} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ic} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rj} & \cdots & a_{rc} \end{bmatrix}$$

  or in a compact form:

$$\mathbf{A} = [a_{ij}], \quad i = 1, \cdots r; j = 1, \cdots, c.$$

- A matrix is said to be **square** if its number of rows equals to the number of columns. Examples:

$$\begin{bmatrix} 4 & 7 \\ 3 & 9 \end{bmatrix}, \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

  Question: what are the dimensions of these matrices? What are their diagonal elements?

- **Vectors**.
  - A matrix containing only one column is called a **column vector**. Example:

$$\mathbf{A} = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix}$$

  - A matrix containing only one row is called a **row vector**. Example:

$$\mathbf{B}' = \begin{bmatrix} 4 & 7 & 10 \end{bmatrix}$$

    $\mathbf{A}$ and $\mathbf{B}'$ are not the same matrix! **For matrices, not only the elements, but also how these elements are arranged matters.**
  - By default, when we say a vector it means a column vector. We also use the prime symbol "$'$" to denote row vectors.

- **Transpose**: the transpose of a matrix $\mathbf{A}$ is another matrix, denoted by $\mathbf{A}'$, that is obtained by interchanging corresponding columns and rows of matrix $\mathbf{A}$.
  - Example: a 3×2 matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 7 & 10 \\ 3 & 4 \end{bmatrix}$$

    Its transpose is a 2×3 matrix

$$\mathbf{A}' = \begin{bmatrix} 2 & 7 & 3 \\ 5 & 10 & 4 \end{bmatrix}$$

    The first row of $\mathbf{A}'$ is the first column of $\mathbf{A}$, the second row of $\mathbf{A}'$ is the second column of $\mathbf{A}$.
  - Example: a column vector

$$\mathbf{C} = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix}$$

    Its transpose is a row vector

$$\mathbf{C}' = \begin{bmatrix} 4 & 7 & 10 \end{bmatrix}$$

    The transpose of a column vector is a row vector.

- In general, if $\mathbf{A}$ is a $r \times c$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1c} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2c} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{ic} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rj} & \cdots & a_{rc} \end{bmatrix} = [a_{ij}] \quad i = 1, \cdots r, j = 1, \cdots c$$

  then its transpose is a $c \times r$ matrix

$$\mathbf{A}' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{i1} & \cdots & a_{r1} \\ a_{12} & a_{22} & \cdots & a_{i2} & \cdots & a_{r2} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{1j} & a_{2j} & \cdots & a_{ij} & \cdots & a_{rj} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{1c} & a_{2c} & \cdots & a_{ic} & \cdots & a_{rc} \end{bmatrix} = [a_{ji}] \quad j = 1, \cdots c, i = 1 \cdots r.$$

  The element in the *ith* row and the *jth* column in A is found in the *jth* row and *ith* column of A′.

- **Equality of matrices**: two matrices are equal if and only if (i) they have the same dimension and (ii) all the corresponding elements are equal.
  - $\mathbf{A} = (\mathbf{A}')'$: a matrix equals to the transpose of its transpose.
- Simple linear regression.
  - In simple linear regression, one basic matrix is the **response vector Y**: an $n \times 1$ column vector

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

    The transpose of $\mathbf{Y}$ is

$$\mathbf{Y}' = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_n \end{bmatrix}$$

    a row vector.
  - Another basic matrix is the **design matrix X**: an $n \times 2$ matrix

$$\mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$\mathbf{X}$ consists of a column of 1s and a column of the observations on the predictor variable X. The transpose of $\mathbf{X}$ is a $2 \times n$ matrix:

$$\mathbf{X}' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \end{bmatrix}$$

2 Matrix algebra

- **Matrix addition and subtraction**.
  - The two matrices must have the same dimension.
  - Their sum (or difference) is another matrix (of the same dimension) whose elements are the sum (or difference) of the corresponding elements of the two matrices.
  - Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

    Then

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{bmatrix}$$

    Question: what is $\mathbf{A} - \mathbf{B}$?

  - In general: if

$$\mathbf{A} = [a_{ij}], \quad \mathbf{B} = [b_{ij}], \quad i = 1, \cdots r, j = 1, \cdots, c$$

    then

$$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}], \quad \mathbf{A} - \mathbf{B} = [a_{ij} - b_{ij}].$$

- **Multiplication of a matrix by a scalar**.
  - A **scalar** is just a number.
  - Multiplying a matrix by a scalar results in another matrix (of the same dimension) whose elements are the corresponding elements of the matrix multiplied by the scalar.
  - Example: if

$$\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix}$$

    then

$$4\mathbf{A} = 4 \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 4 \times 2 & 4 \times 7 \\ 4 \times 9 & 4 \times 3 \end{bmatrix} = \begin{bmatrix} 8 & 28 \\ 36 & 12 \end{bmatrix}$$

– In general: if $k \in \mathbb{R}$ is a scalar and

$$\mathbf{A} = [a_{ij}], \quad i = 1, \cdots r, j = 1, \cdots, c,$$

then

$$k\mathbf{A} = [ka_{ij}].$$

- **Multiplication of a matrix by a matrix**.

  – **Matrix multiplication is only defined when the two matrices have conforming dimensions** which means: the number of columns of the first matrix (the matrix on left) equals to the number of rows of the second matrix (the matrix on right).

  – Example:

$$\mathbf{A} = \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

  Then **AB** is well defined, whereas **AC**, **BC** are not defined. Questions, how about **BA**?

  – In the above example, the product **AB** will be a 2×1 matrix. Its $(1, 1)$ element is the sum of cross products of the $1st$ row of matrix **A** with the $1st$ column of matrix **B**. Its $(2, 1)$ element is the sum of cross products of the $2nd$ row of matrix **A** with the $1st$ column of matrix **B**.

$$\mathbf{AB} = \begin{bmatrix} 2 & 7 \\ 9 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \times 1 + 7 \times 2 \\ 9 \times 1 + 3 \times 2 \end{bmatrix} = \begin{bmatrix} 16 \\ 15 \end{bmatrix}$$

  – In general: if $\mathbf{A} = [a_{ij}]$ is a $r \times c$ matrix and $\mathbf{B} = [b_{ij}]$ is a $c \times s$ matrix, then the product **AB** is a $r \times s$ matrix whose element in the ith row and jth column is

$$\sum_{k=1}^{c} a_{ik} b_{kj}, \quad i = 1, \cdots r, j = 1, \cdots s.$$

- Example.

$$\begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} = \begin{bmatrix} 2 \times 4 + 5 \times 5 & 2 \times 6 + 5 \times 8 \\ 4 \times 4 + 1 \times 5 & 4 \times 6 + 1 \times 8 \end{bmatrix} = \begin{bmatrix} 33 & 52 \\ 21 & 32 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 6 \\ 5 & 8 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 4 \times 2 + 6 \times 4 & 4 \times 5 + 6 \times 1 \\ 5 \times 2 + 8 \times 4 & 5 \times 5 + 8 \times 1 \end{bmatrix} = \begin{bmatrix} 32 & 26 \\ 42 & 33 \end{bmatrix}$$

  Question: do you notice something?

- In matrix algebra, usually $\mathbf{AB} \neq \mathbf{BA}$. So **the order of multiplication does matter.** Indeed, even when the product **AB** is well defined, the product **BA** may not be defined at all.

- Example: if

$$\mathbf{A} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is a 3×1 column vector, then

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \times 1 + 2 \times 2 + 3 \times 3 = 14$$

is a 1×1 matrix, i.e., a scalar. And

$$\mathbf{A}\mathbf{A}' = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

is a 3×3 matrix.

- Simple linear regression. The regression equations:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad i = 1, \cdots n$$

can be written in a compact matrix form.

  - The **response vector Y**: an $n \times 1$ column vector

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_i \\ \vdots \\ Y_n \end{bmatrix}$$

  - The **error vector** : an $n \times 1$ column vector

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_n \end{bmatrix}$$

  - The **coefficient vector**: a 2×1 column vector:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

– The **design matrix**: an $n \times 2$ matrix

$$\mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_i \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

– The product of $\mathbf{X}$ and $\boldsymbol{\beta}$

$$\mathbf{X}\boldsymbol{\beta} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_i \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_1 \\ \beta_0 + \beta_1 X_2 \\ \vdots \\ \beta_0 + \beta_1 X_i \\ \vdots \\ \beta_0 + \beta_1 X_n \end{bmatrix} = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_i) \\ \vdots \\ E(Y_n) \end{bmatrix} = \mathbf{E}(\mathbf{Y})$$

is an $n \times 1$ column vector – the **mean vector**.

– **The matrix form of simple linear regression model equations**:

$$\underset{n\times 1}{\mathbf{Y}} = \underset{n\times 2}{\mathbf{X}} \underset{2\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\epsilon}}.$$

3 Matrix algebra (Cont'd)

- Special matrices.

  – **Vector of Zeros**: an $r \times 1$ vector with all elements being zero

$$\mathbf{0}_r = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

  Example:

$$\mathbf{0}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

  – **Vector of ones**: an $r \times 1$ vector with all elements being one

$$\mathbf{1}_r = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Example:

$$\mathbf{1}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

– **Matrix of ones**: an $r \times r$ matrix with all elements being one

$$\boldsymbol{J}_r = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ldots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}.$$

$$\boldsymbol{J}_r = \mathbf{1}_r \mathbf{1}_r'.$$

Example:

$$\boldsymbol{J}_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

– **Identity matrix**: an $r \times r$ square matrix with all diagonal elements being 1 and all other elements being 0

$$\mathbf{I}_r = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Example:

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

– **Diagonal matrix**: an $r \times r$ square matrix with all off-diagonal elements being 0

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_r \end{bmatrix}$$

This matrix can be denoted by $diag(d_1, \cdots, d_r)$. Note that some of d_is may be zero.
Example:

$$\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\mathbf{I}_r$ is a diagonal matrix with all diagonal elements being one.

- **Symmetric matrices**: an $r \times r$ (square) matrix $\mathbf{H}$ is symmetric if it equals to its transpose:

$$\mathbf{H}' = \mathbf{H}.$$

This is equivalent to saying that its (i,j) element equals to its (j,i) element for all $1 \le i,j \le r$.

  * Identity matrices and diagonal matrices are symmetric.
  * Any matrix in the form $\mathbf{C}'\mathbf{C}$ is symmetric. Question: why?

- **Linear dependence of a set of vectors**.

  - Let $\mathbf{C}_1, \cdots, \mathbf{C}_s$ be $r \times 1$ column vectors. They are said to be **linearly dependent** if and only if there exist constants $k_1, \cdots, k_s$, **not all zero**, such that

$$k_1\mathbf{C}_1 + \cdots + k_s\mathbf{C}_s = \mathbf{0},$$

where $\mathbf{0}$ is the $r \times 1$ vector with all elements being zero. This is equivalent to saying that one of these vectors can be written as a linear combination of the rest of the vectors.

  - Example.

$$\mathbf{C}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix}2\\2\\4\end{bmatrix}, \quad \mathbf{C}_3 = \begin{bmatrix}4\\6\\10\end{bmatrix}$$

Note that, $2\mathbf{C}_1 + \mathbf{C}_2 = \mathbf{C}_3$, so $2\mathbf{C}_1 + \mathbf{C}_2 + (-1)\mathbf{C}_3 = \mathbf{0}$. Thus, $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$ are linearly dependent.

  - **If the number of vectors s is greater than the dimension of the vectors r, then $\mathbf{C}_1, \cdots, \mathbf{C}_s$ must be linearly dependent.** For example:

$$\mathbf{C}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \quad \mathbf{C}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}, \quad \mathbf{C}_4 = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

are linearly dependent. Question: are you able to express one of C_is as a linear combination of the others in the above example?

  - If $\mathbf{C}_1, \cdots, \mathbf{C}_s$ are linearly dependent, then adding any $r \times 1$ vector $\mathbf{C}$ into the set still makes a linearly dependent set of vectors. Question: why?

- **Linear independence of a set of vectors**.

  - Let $\mathbf{C}_1, \cdots, \mathbf{C}_s$ be $r \times 1$ column vectors. If they are not linearly dependent, then they are said to be **linearly independent**.
  - $\mathbf{C}_1, \cdots, \mathbf{C}_s$ are linearly independent if and only if

$$k_1\mathbf{C}_1 + \cdots + k_s\mathbf{C}_s = \mathbf{0}$$

implies $k_1 = \cdots = k_s = 0$.

- A set with a single **nonzero vector** must be linearly independent. Question: how about a single zero vector?
- Example.

$$\mathbf{C}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{C}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{C}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

  are linearly independent. Question: why?

- **Rank of a matrix**.
  - **The rank of a matrix is the maximum number of linearly independent columns in the matrix.**
  - Example.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

    Use $\mathbf{C}_j$ to denote the jth column of matrix $\mathbf{A}$ ($j = 1, \cdots 4$). We know that $\mathbf{C}_1, \cdots, \mathbf{C}_4$ are linearly dependent, but $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$ are linearly independent. So $rank(\mathbf{A}) = 3$.
  - The rank of a matrix is unique. It also equals to the maximum number of linearly independent rows in the matrix.
  - For an $r \times s$ matrix $\mathbf{A}$, $rank(\mathbf{A}) \leq \min\{r, s\}$. That is, the rank of a matrix can not exceed either its row dimension or its column dimension.
  - If $\mathbf{C} = \mathbf{AB}$, then $rank(\mathbf{C}) \leq \min\{rank(\mathbf{A}), rank(\mathbf{B})\}$. That is, the rank of the product matrix can not exceed the rank of either of the matrices in the product.
  - Examples.

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

    $rank(\mathbf{I}_3) = 3, rank(\mathbf{D}) = 2$.
  - The rank of an identity matrix equals to its dimension: $rank(\mathbf{I}_r) = r$. The rank of a diagonal matrix equals to its number of nonzero diagonal elements. Question: what is the rank of $\mathbf{0}_r$? and what is the rank of $\boldsymbol{J}_r$?

- **Matrix inversion**.
  - If an $r \times r$ (square) matrix $\mathbf{A}$ has **full rank**, i.e., $rank(\mathbf{A}) = r$, then $\mathbf{A}$ is said to be **nonsingular**. Otherwise if $rank(\mathbf{A}) < r$, $\mathbf{A}$ is said to be **singular**.
    * $\mathbf{A}$ being nonsingular is equivalent to saying that all columns (equiv. rows) of $\mathbf{A}$ are linearly independent.

* Example.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

  A is nonsingular, whereas **B** is singular. Question: why?

* Identity matrices are nonsingular. A diagonal matrix is nonsingular if and only if all its diagonal elements are nonzero.

- If an $r \times r$ matrix $\mathbf{A}$ is nonsingular, then there exists a unique $r \times r$ matrix, denoted by $\mathbf{A}^{-1}$ such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_r,$$

  where $\mathbf{I}_r$ is the $r \times r$ identity matrix. $\mathbf{A}^{-1}$ is called the **inverse** of matrix $\mathbf{A}$. It also follows that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_r.$$

- Inverse is only defined for square matrices. Even so, many square matrices do not have inverses (if they are singular).
- The inverse of an identity matrix is itself: $(\mathbf{I}_r)^{-1} = \mathbf{I}_r$.
- The inverse of a diagonal matrix (if exists) is a diagonal matrix consisting of the reciprocals of the elements on the diagonal. If

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_r \end{bmatrix}$$

  where $d_1, \cdots, d_r$ are nonzero, then

$$\mathbf{D}^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_r} \end{bmatrix}$$

- **Inverse of 2×2 matrices.**

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

  - $\mathbf{A}$ is nonsingular (hence invertible) if and only if

$$ad - bc \neq 0.$$

  - When $ad - bc \neq 0$, the inverse of $\mathbf{A}$ is [check this out!]

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

– Examples.

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 0.5 & 1 \end{bmatrix}$$

Since

$$2 \times 1 - 4 \times 3 = -10 \neq 0$$

$\mathbf{A}$ is invertible,

$$\mathbf{A}^{-1} = \frac{1}{-10} \begin{bmatrix} 1 & -4 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} -0.1 & 0.4 \\ 0.3 & -0.2 \end{bmatrix}$$

As for $\mathbf{B}$, since $2 \times 1 - 4 \times 0.5 = 0$, $\mathbf{B}$ is not invertible. Questions: what are $rank(\mathbf{A})$ and $rank(\mathbf{B})$?

- Simple linear regression: we need invert the matrix $\mathbf{X}'\mathbf{X}$.

  – The **design matrix**: an $n \times 2$ matrix

$$\mathbf{X} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_i \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_i & \cdots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_i \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{bmatrix}$$

is a 2×2 matrix.

  – $a = n, \quad b = c = \sum_{i=1}^n X_i, \quad d = \sum_{i=1}^n X_i^2$.

$$D = ad - bc = n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 = n \sum_{i=1}^n (X_i - \overline{X})^2.$$

When $D \neq 0$,

$$\begin{aligned} \left(\mathbf{X}'\mathbf{X}\right)^{-1} &= \begin{bmatrix} \frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n (X_i - \overline{X})^2} & -\frac{\sum_{i=1}^n X_i}{n \sum_{i=1}^n (X_i - \overline{X})^2} \\ -\frac{\sum_{i=1}^n X_i}{n \sum_{i=1}^n (X_i - \overline{X})^2} & \frac{n}{n \sum_{i=1}^n (X_i - \overline{X})^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i - \overline{X})^2} & -\frac{\overline{X}}{\sum_{i=1}^n (X_i - \overline{X})^2} \\ -\frac{\overline{X}}{\sum_{i=1}^n (X_i - \overline{X})^2} & \frac{1}{\sum_{i=1}^n (X_i - \overline{X})^2} \end{bmatrix} \end{aligned}$$

- Notes: $D = n\sum_{i=1}^{n}(X_i - \overline{X})^2 \neq 0$ if and only if X_is are not all equal.

- For general $r \times r$ matrices, we often use computers to check whether it is invertible and to find its inverse (if exists).

4 Basic facts about matrix algebra

- $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$; $(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$
- $(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})$; Note often, $\mathbf{AB}\neq\mathbf{BA}$
- $\mathbf{C}(\mathbf{A}+\mathbf{B})=\mathbf{CA}+\mathbf{CB}$
- $k(\mathbf{A}+\mathbf{B})=k\mathbf{A}+k\mathbf{B}$
- $(\mathbf{A}')'=\mathbf{A}$; $(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}'$; $(\mathbf{AB})'=\mathbf{B}'\mathbf{A}'$ (note the interchange of order here!)
- $((\mathbf{A})^{-1})^{-1}=\mathbf{A}$; $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$ (note the interchange of order here!)
- $(\mathbf{A}')^{-1}=(\mathbf{A}^{-1})'$
- For an $r\times s$ matrix $\mathbf{A}$, $rank(\mathbf{A})\leq\min\{r,s\}$; If $\mathbf{C}=\mathbf{AB}$, then $rank(\mathbf{C})\leq\min\{rank(\mathbf{A}),rank(\mathbf{B})\}$
- The inverse of an identity matrix is itself: $(\mathbf{I}_r)^{-1}=\mathbf{I}_r$.
- The inverse of a diagonal matrix exists if and only if all its diagonal elements are nonzero and the inverse is a diagonal matrix consisting of the reciprocals of the elements on the diagonal.
- If $\mathbf{B}$ is an $r\times s$ matrix and $\mathbf{I}_r$ is the $r\times r$ identity matrix, then

$$\mathbf{I}_r\mathbf{B}=\mathbf{B}.$$

- If $\mathbf{B}$ is an $r\times r$ square matrix and $\mathbf{I}_r$ is the $r\times r$ identity matrix, then

$$\mathbf{I}_r\mathbf{B}=\mathbf{B}\mathbf{I}_r=\mathbf{B}.$$

5 Random Vectors

- A **random vector (or matrix)** contains elements that are random variables.
- **Expectation of a random vector (or matrix).**
  - A vector (or matrix) of the same dimension whose elements are the expected values of the respective random variables.
  - Suppose $\mathbf{Z}=[Z_{ij}]$ is an $r\times s$ random matrix, then its expectation is an $r\times s$ matrix defined as:

$$\mathbf{E}\{\mathbf{Z}\}:=[E(Z_{ij})],\quad i=1,\cdots,r, j=1,\cdots,s.$$

- **Variance-Covariance matrix of a random vector**.
  - Then variance-Covariance matrix of a random vector is a square matrix whose diagonal elements are the variance of the respective random variables and whose off-diagonal elements are the covariance between respective random variables.
  - Suppose $\mathbf{Z}$ is an $r \times 1$ random vector:

$$\mathbf{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_r \end{bmatrix}$$

    Then its variance-covariance matrix is an $r \times r$ matrix defined as:

$$\boldsymbol{\sigma}^2\{\mathbf{Z}\} := \begin{bmatrix} Var(Z_1) & Cov(Z_1, Z_2) & \cdots & Cov(Z_1, Z_r) \\ Cov(Z_2, Z_1) & Var(Z_2) & \cdots & Cov(Z_2, Z_r) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(Z_r, Z_1) & Cov(Z_r, Z_2) & \cdots & Var(Z_r) \end{bmatrix}$$

  - Note $Var(Z_i)$ can also be written as $Cov(Z_i, Z_i)$. So the (i, j)th element of $\boldsymbol{\sigma}^2\{\mathbf{Z}\}$ is $Cov(Z_i, Z_j)$.
  - $\boldsymbol{\sigma}^2\{\mathbf{Z}\}$ is symmetric because

$$Cov(Z_i, Z_j) = Cov(Z_j, Z_i).$$

  - Since $Cov(Z_i, Z_j) = E\left\{(Z_i - E(Z_i))(Z_j - E(Z_j))\right\}$, it follows that:

$$\boldsymbol{\sigma}^2\{\mathbf{Z}\} = \mathbf{E}\{(\mathbf{Z} - \mathbf{E}(\mathbf{Z}))\,(\mathbf{Z} - \mathbf{E}(\mathbf{Z}))'\}.$$

- **Linear transformations of a random vector.**
  - $\mathbf{Z}$ is an $r \times 1$ random vector, and $\mathbf{A}$ is an $s \times r$ non-random matrix. Then

$$\underset{s\times 1}{\mathbf{W}} = \underset{s\times r}{\mathbf{A}}\;\underset{r\times 1}{\mathbf{Z}}$$

    is an $s \times 1$ random vector.
  - Properties:

$$\begin{aligned} \mathbf{E}\{\mathbf{A}\} &= \mathbf{A}. \\ \mathbf{E}\{\mathbf{W}\} &= \mathbf{E}\{\mathbf{AZ}\} = \mathbf{AE}\{\mathbf{Z}\}. \\ \boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{W}\} &= \boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{AZ}\} = \mathbf{A}\boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{Z}\}\mathbf{A}'. \end{aligned}$$

  - Special case: $\mathbf{Z}$ is an $r \times 1$ random vector, and $\mathbf{C}$ is an $r \times 1$ non-random vector. Then

$$U = \mathbf{C}'\mathbf{Z}.$$

    is a random variable with

$$E(U) = \mathbf{C}'\mathbf{E}\{\mathbf{Z}\}.$$
$$Var(U) = \mathbf{C}'\boldsymbol{\sigma}^{\mathbf{2}}\{\mathbf{Z}\}\mathbf{C}.$$

- **Quadratic forms.**
  - A quadratic form of an $n \times 1$ vector $\mathbf{Y}$ is of the form:

    $$\mathbf{Y}'\mathbf{A}\mathbf{Y},$$

    where $\mathbf{A} = [a_{ij}]$ is an $n \times n$ **symmetric matrix**, i.e., $a_{ij} = a_{ji}$ for all $1 \leq i, j \leq n$.
  - $\mathbf{Y}'\mathbf{A}\mathbf{Y} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} Y_i Y_j$.
  - Quadratic forms play an important role in statistics because all sums of squares in ANOVA for linear models can be expressed as quadratic forms.
  - SSTO, SSE, SSR are quadratic forms in the response vector $\mathbf{Y}$.

6 Exercises

- Tell true or false of the following statements. Briefly justify your answer.

  (a)
  $$\mathbf{A} = \begin{bmatrix} & 14 & \\ 10 & & 15 \\ 1 & 2 & 3 \end{bmatrix}$$
  is **not** a matrix.

  (b)
  $$\begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 4 & 7 & 10 \end{bmatrix}$$
  are the same matrix.

  (c) The transpose of a column vector is a row vector.

  (d) If both $\mathbf{A}$ and $\mathbf{B}$ are 3×4 matrices, then $\mathbf{AB}$ is **not** defined.

  (e) In matrix multiplication, usually $\mathbf{AB} \neq \mathbf{BA}$.

  (f) $\mathbf{I}_3$ is the 3×3 identity matrix. Then for **any** 3×3 matrix $\mathbf{A}$, we have
  $$\mathbf{I}_3\mathbf{A} = \mathbf{A}\mathbf{I}_3 = \mathbf{A}.$$

  (g) Diagonal matrices are always invertible.

- **Inverse of** 2×2 **matrices**. Suppose
  $$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
  and $ad - bc \neq 0$. Let
  $$\mathbf{B} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
  Show that
  $$\mathbf{BA} = \mathbf{AB} = \mathbf{I}_2.$$
  Therefore, by definition $\mathbf{B}$ is the inverse of $\mathbf{A}$, i.e., $\mathbf{B} = \mathbf{A}^{-1}$.

  (Recall: $\mathbf{I}_2$ is the 2×2 identity matrix:
  $$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
)

- **Matrix algebra.**
  $$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 6 \\ 9 \\ 1 \end{bmatrix}$$

(a) State the dimension of each matrix. Between which matrices addition and substraction are allowed? Between which matrices multiplication is allowed?

(b) What is the $(1, 2)$- element of $\mathbf{A}$? What are the diagonal elements of $\mathbf{A}$? What is $\mathbf{A}'$? What is $(\mathbf{A}')'$?

(c) Calculate the following:

$$\mathbf{A}+\mathbf{B}, \quad \mathbf{A}-\mathbf{B}, \quad 4\mathbf{A}, \quad \mathbf{AC}, \quad \mathbf{B}'\mathbf{A}, \quad \mathbf{DD}', \quad \mathbf{D}'\mathbf{D}$$

and state the dimensions of the resulting matrices. Which of these matrices are symmetric?

(d) Is $\mathbf{B}'\mathbf{A}$ invertible? If so, find its inverse.

(e) What are the ranks of $\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{B}'\mathbf{A}, \mathbf{DD}'$?

- Consider the simultaneous equations:

$$\begin{aligned} 5y_1 + 2y_2 &= 8 \\ 23y_1 + 7y_2 &= 28 \end{aligned}$$

(a) Write these equations in matrix form.

(b) Use matrix algebra to find the solutions.

- **Random vector**. Consider the following linear transformations of the random variables Y_1, Y_2 and Y_3:

$$\begin{aligned} W_1 &= Y_1 + Y_2 + Y_3 \\ W_2 &= Y_1 - Y_2 \\ W_3 &= Y_1 - Y_2 - Y_3. \end{aligned}$$

(a) State the above in matrix form.

(b) Let $\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$ and $\mathbf{W} = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix}$. If the expectation of $\mathbf{Y}$ is

$$\mathbf{E}\{\mathbf{Y}\} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

find $\mathbf{E}\{\mathbf{W}\}$.

(c) If the variance-covariance matrix of $\mathbf{Y}$ is

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\} = \begin{bmatrix} 1 & 0.5 & 0.8 \\ 0.5 & 1 & 0.7 \\ 0.8 & 0.7 & 1 \end{bmatrix},$$

find the variance-covariance matrix of $\mathbf{W}$.

7 Solution to exercises

- Tell true or false of the following statements. Briefly justify your answer.

  (a)
  $$\mathbf{A} = \begin{bmatrix} & 14 & \\ 10 & & 15 \\ 1 & 2 & 3 \end{bmatrix}$$
  is **not** a matrix.

  True. Not rectangular.

  (b)
  $$\begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 4 & 7 & 10 \end{bmatrix}$$
  are the same matrix.

  False. One is 3×1(column vector) and the other is 1×3(row vector).

  (c) The transpose of a column vector is a row vector.

  True.

  (d) If both $\mathbf{A}$ and $\mathbf{B}$ are 3×4 matrices, then $\mathbf{AB}$ is **not** defined.

  True. the number of columns of the first matrix (the matrix on left) doesn't equal the number of rows of the second matrix (the matrix on right).

  (e) In matrix multiplication, usually $\mathbf{AB} \neq \mathbf{BA}$.

  True.

  (f) $\mathbf{I}_3$ is the 3×3 identity matrix. Then for **any** 3×3 matrix $\mathbf{A}$, we have
  $$\mathbf{I}_3\mathbf{A} = \mathbf{A}\mathbf{I}_3 = \mathbf{A}.$$

  True.

  (g) Diagonal matrices are always invertible.

  False. The inverse of a diagonal matrix exists if and only if all its diagonal elements are nonzero.

- **Inverse of 2×2 matrices**. Suppose
  $$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
  and $ad - bc \neq 0$. Let
  $$\mathbf{B} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
  Show that
  $$\mathbf{BA} = \mathbf{AB} = \mathbf{I}_2.$$

Therefore, by definition $\mathbf{B}$ is the inverse of $\mathbf{A}$, i.e., $\mathbf{B} = \mathbf{A}^{-1}$.

(Recall: $\mathbf{I}_2$ is the 2×2 identity matrix:

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

)

$$\mathbf{BA} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & bd-bd \\ -ac+ac & -bc+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2$$

$$\mathbf{AB} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ab \\ cd-cd & -bc+ad \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2$$

- **Matrix algebra.**

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 6 \\ 9 \\ 1 \end{bmatrix}$$

(a) State the dimension of each matrix. Between which matrices addition and substraction are allowed? Between which matrices multiplication is allowed?

$\mathbf{A}$ and $\mathbf{B}$ is 3×2, $\mathbf{C}$ is 2×3 and $\mathbf{D}$ is 3×1. Addition and subtraction are allowed between $\mathbf{A}$ and $\mathbf{B}$. Multiplication is allowed for $\mathbf{AC}$, $\mathbf{CA}$, $\mathbf{BC}$, $\mathbf{CB}$ and $\mathbf{CD}$.

(b) What is the $(1,2)$- element of $\mathbf{A}$? What are the diagonal elements of $\mathbf{A}$? What is $\mathbf{A}'$? What is $(\mathbf{A}')'$?

The $(1,2)$-element of $\mathbf{A}$ is 4. The diagonal elements of $\mathbf{A}$ are 1 and 6. $\mathbf{A}' = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix}' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 6 & 8 \end{bmatrix}$. $(\mathbf{A}')' = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 6 & 8 \end{bmatrix}' = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} = \mathbf{A}$.

(c) Calculate the following:

$$\mathbf{A}+\mathbf{B}, \quad \mathbf{A}-\mathbf{B}, \quad 4\mathbf{A}, \quad \mathbf{AC}, \quad \mathbf{B}'\mathbf{A}, \quad \mathbf{DD}', \quad \mathbf{D}'\mathbf{D}$$

and state the dimensions of the resulting matrices. Which of these matrices are symmetric?

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1+1 & 4+3 \\ 2+1 & 6+4 \\ 3+2 & 8+5 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ 3 & 10 \\ 5 & 13 \end{bmatrix}_{3\times 2}$$

$$\mathbf{A}-\mathbf{B} = \begin{bmatrix} 1 & 4 \\ 2 & 6 \\ 3 & 8 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1-1 & 4-3 \\ 2-1 & 6-4 \\ 3-2 & 8-5 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}_{3\times 2}$$

$$4\mathbf{A} = 4\begin{bmatrix}1 & 4\\2 & 6\\3 & 8\end{bmatrix} = \begin{bmatrix}4\times 1 & 4\times 4\\4\times 2 & 4\times 6\\4\times 3 & 4\times 8\end{bmatrix} = \begin{bmatrix}4 & 16\\8 & 24\\12 & 32\end{bmatrix}_{3\times 2}$$

$$\mathbf{AC} = \begin{bmatrix}1 & 4\\2 & 6\\3 & 8\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\2 & 4 & 6\end{bmatrix} = \begin{bmatrix}1\times 1+4\times 2 & 1\times 2+4\times 4 & 1\times 3+4\times 6\\2\times 1+6\times 2 & 2\times 2+6\times 4 & 2\times 3+6\times 6\\3\times 1+8\times 2 & 3\times 2+8\times 4 & 3\times 3+8\times 6\end{bmatrix} = \begin{bmatrix}9 & 18 & 27\\14 & 28 & 42\\19 & 38 & 57\end{bmatrix}_{3\times 3}$$

$$\mathbf{B}'\mathbf{A} = \begin{bmatrix}1 & 1 & 2\\3 & 4 & 5\end{bmatrix}\begin{bmatrix}1 & 4\\2 & 6\\3 & 8\end{bmatrix} = \begin{bmatrix}1\times 1+1\times 2+2\times 3 & 1\times 4+1\times 6+2\times 8\\3\times 1+4\times 2+5\times 3 & 3\times 4+4\times 6+5\times 8\end{bmatrix} = \begin{bmatrix}9 & 26\\26 & 76\end{bmatrix}_{2\times 2}$$

$$\mathbf{DD}' = \begin{bmatrix}6\\9\\1\end{bmatrix}\begin{bmatrix}6 & 9 & 1\end{bmatrix} = \begin{bmatrix}6\times 6 & 6\times 9 & 6\times 1\\9\times 6 & 9\times 9 & 9\times 1\\1\times 6 & 1\times 9 & 1\times 1\end{bmatrix} = \begin{bmatrix}36 & 54 & 6\\54 & 81 & 9\\6 & 9 & 1\end{bmatrix}_{3\times 3}$$

$$\mathbf{D}'\mathbf{D} = \begin{bmatrix}6 & 9 & 1\end{bmatrix}\begin{bmatrix}6\\9\\1\end{bmatrix} = 6\times 6+9\times 9+1\times 1 = 118_{1\times 1}$$

$\mathbf{B}'\mathbf{A}$, $\mathbf{DD}'$, and $\mathbf{D}'\mathbf{D}$ are symmetric.

(d) Is $\mathbf{B}'\mathbf{A}$ invertible? If so, find its inverse.
Since $9\times 76 - 26\times 26 = 8 \neq 0$, $\mathbf{B}'\mathbf{A}$ is invertible.

$$(\mathbf{B}'\mathbf{A})^{-1} = \frac{1}{9\times 76 - 26\times 26}\begin{bmatrix}76 & -26\\-26 & 9\end{bmatrix} = \begin{bmatrix}9.5 & -3.25\\-3.25 & 1.125\end{bmatrix}$$

(e) What are the ranks of $\mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{B}'\mathbf{A}, \mathbf{DD}'$?
$rank(\mathbf{B}) = 2$, $rank(\mathbf{C}) = 1$, $rank(\mathbf{D}) = 1$, $rank(\mathbf{B}'\mathbf{A}) = 2$, $rank(\mathbf{DD}') = 1$

- Consider the simultaneous equations:

$$\begin{aligned}5y_1 + 2y_2 &= 8\\23y_1 + 7y_2 &= 28\end{aligned}$$

(a) Write these equations in matrix form.

$$\begin{bmatrix}5 & 2\\23 & 7\end{bmatrix}\begin{bmatrix}y_1\\y_2\end{bmatrix} = \begin{bmatrix}8\\28\end{bmatrix}$$

(b) Use matrix algebra to find the solutions.

$$\begin{aligned}\begin{bmatrix}y_1\\y_2\end{bmatrix} &= \begin{bmatrix}5 & 2\\23 & 7\end{bmatrix}^{-1}\begin{bmatrix}8\\28\end{bmatrix}\\ &= \frac{1}{5\times 7 - 23\times 2}\begin{bmatrix}7 & -2\\-23 & 5\end{bmatrix}\begin{bmatrix}8\\28\end{bmatrix}\\ &= -\frac{1}{11}\begin{bmatrix}0\\-44\end{bmatrix}\\ &= \begin{bmatrix}0\\4\end{bmatrix}\end{aligned}$$

- **Random vector.** Consider the following linear transformations of the random variables Y_1, Y_2 and Y_3:

$$\begin{aligned} W_1 &= Y_1 + Y_2 + Y_3 \\ W_2 &= Y_1 - Y_2 \\ W_3 &= Y_1 - Y_2 - Y_3. \end{aligned}$$

(a) State the above in matrix form.

$$\begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$$

(b) Let $\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$ and $\mathbf{W} = \begin{bmatrix} W_1 \\ W_2 \\ W_3 \end{bmatrix}$. If the expectation of $\mathbf{Y}$ is

$$\mathbf{E}\{\mathbf{Y}\} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

find $\mathbf{E}\{\mathbf{W}\}$.

Let $\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix}$, then $\mathbf{W} = \mathbf{AY}$ and

$$\mathbf{E}\{\mathbf{W}\} = \mathbf{E}\{\mathbf{AY}\} = \mathbf{AE}\{\mathbf{Y}\} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -1 \\ -4 \end{bmatrix}$$

(c) If the variance-covariance matrix of $\mathbf{Y}$ is

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\} = \begin{bmatrix} 1 & 0.5 & 0.8 \\ 0.5 & 1 & 0.7 \\ 0.8 & 0.7 & 1 \end{bmatrix},$$

find the variance-covariance matrix of $\mathbf{W}$.

$$\begin{aligned} \boldsymbol{\sigma}^2\{\mathbf{W}\} &= \boldsymbol{\sigma}^2\{\mathbf{AY}\} \\ &= \mathbf{A}\boldsymbol{\sigma}^2\{\mathbf{Y}\}\mathbf{A}' \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0.5 & 0.8 \\ 0.5 & 1 & 0.7 \\ 0.8 & 0.7 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & 0.1 & -2.4 \\ 0.1 & 1 & 0.9 \\ -2.4 & 0.9 & 1.8 \end{bmatrix} \end{aligned}$$