Lecture Note 01: *Linear Algebra*

March 15, 2022

Introduction

tbw

Contents

List of Defintions

List of Theorems and Propositions

List of Examples

1 Terminology

Definition 1.1 Matrix

A **matrix** is a rectangular array of numbers arranged in *rows* (**n**) and *columns* (**K**):

$$\mathbf{A_{n,K}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix}$$

An element $a_{i,k} \in \mathbf{A}$ represents the element positioned in row i and column k.

Definition 1.2 Vector

A **vector** is an ordered set of numbers arranged in either a single row or a single column.

A **row vector** (**r**) is a vector with just one row:

$$r_{1,K} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \end{bmatrix}$$

A **column vector** (**c**) is a vector with just one column.

$$c_{n,1} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}$$

A matrix can be viewed as a set of vectors. The **dimension** of a matrix is the number of rows (**n**) and columns (**K**) it contains. Hence, a matrix **A** with n rows and K columns has a dimension of $n \times K$.

Definition 1.3 Scalar

A **scalar** is a single number, scalars are simultaneously column and row vectors.

The **main diagonal** of a matrix is the set of elements $a_{i,k}$ of a matrix such that $i = k$. With this brief terminology we can define the following:

- **Square matrix**: A matrix with the same number of rows and columns.

Example 1.1 Square matrices

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} ; \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1K} \\ b_{21} & b_{22} & \cdots & b_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nK} \end{bmatrix}, \ where \ \ n = K \ ; \mathbf{C} = \begin{bmatrix} c_{11} \end{bmatrix}$$

- **Symmetric matrix**: A square matrix such that: $a_{i,k} = a_{k,i}$ for all i and k.

Example 1.2 Symmetric matrices

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} ; \mathbf{B} = \begin{bmatrix} 1 & -5 & 3 \\ -5 & -2 & 9 \\ 3 & 9 & 2 \end{bmatrix} ; \mathbf{C} = \begin{bmatrix} c_{11} \end{bmatrix}$$

- **Diagonal matrix**: A square matrix such that: (1) at least one element of the main diagonal is non-zero, and, (2) all element that is not part of the main diagonal is zero. Note that diagonal matrices are also symmetric matrices.

Example 1.3 Diagonal Matrices

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} ; \mathbf{B} = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_{nK} \end{bmatrix}, \ where \ \ n = K \ ; \mathbf{C} = \begin{bmatrix} c_{11} \end{bmatrix}$$

Do not confuse diagonal matrix with diagonal of a matrix. The elements on positions where (number of rows) = (number of columns) like a_{11}, a_{22}, a_{33} and so on, form diagonal of a matrix. Diagonal exists for rectangular matrix also (and the way of finding diagonal elements remains same).

- **Scalar matrix**: Is a diagonal matrix such that all the elements in the main diagonal are the same.

Example 1.4 Scalar Matrices

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} ; \mathbf{B} = \begin{bmatrix} b & 0 & \cdots & 0 \\ 0 & b & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b \end{bmatrix} ; \mathbf{C} = \begin{bmatrix} c_{11} \end{bmatrix}$$

- **Identity or Unity matrix**: A scalar matrix such the elements in the main diagonal are ones. An identity matrix is often denoted by the symbol **I**. Sometimes the dimension of the matrix is also added as a subscript; $\mathbf{I_n}$ indicates an identity matrix of dimension $n \times n$. It is also common to refer to an $\mathbf{I_n}$ matrix as an identity matrix of *size* n.

Example 1.5 Identity Matrices

$$\mathbf{I_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; \mathbf{I_n} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} ; \mathbf{I_1} = \begin{bmatrix} 1 \end{bmatrix}$$

- **Zero matrix or Null matrix**: A matrix whose elements are all zeros. A null matrix is often denoted by **0**. It is also common to use $\mathbf{0_n}$ to refer to a null matrix of *size* n.

Example 1.6 Zero or Null Matrices

$$\mathbf{0_2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} ; \mathbf{0_n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} ; \mathbf{0_1} = \begin{bmatrix} 0 \end{bmatrix}$$

- **Triangular matrix**: Is a matrix that only has zeros either above or below the main diagonal. If the zeros are above the main diagonal is a **lower triangular** matrix, if the zeros are below is an **upper triangular** matrix. Note that the zeros are *above* or *below* and not in the main diagonal.

A triangular matrix is one that is either lower triangular or upper triangular. A diagonal matrix is both upper and lower triangular.

Example 1.7 Triangular matrices

Lower triangular matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} ; \mathbf{L_{n,K}} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nK} \end{bmatrix}$$

Upper triangular matrices:

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} ; \mathbf{U_{n,K}} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1K} \\ 0 & u_{22} & \cdots & u_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nK} \end{bmatrix}$$

- **Conformable matrix**: Is a matrix that has dimensions suitable for a given operation, *e.g*, if matrices $\mathbf{A}$ and $\mathbf{B}$ have the same dimension then we say that they are *conformable for addition*. Another example, if $\mathbf{A}$ is a square matrix then $\mathbf{A}$ is *conformable for inversion*. More about this when discussing the different operations with matrices.

- Discussion/Motivation:
  - How would data for cross section likely to differ from time series data when put in matrix form?
    - Cross Section ($N \geq K$)
    - Time Series ($N \leq K$)
  - In what type of matrix would the following error structures are likely to be represented?
    - HOMOSKEDASTIC errors (Scalar Matrix).
    - HETEROSKEDASTIC errors (Diagnol Matrix).
    - AUTOCORRELATION matrix (Diagnoal Matrix implies no auto-correlation, Many and 'high' numbers in off-diagonal elements imply strong auto-correlation)
  - Does $\mathbf{y} = \beta\mathbf{X}$ imply we can always solve for $\beta^* = \mathbf{X}^{-1}\mathbf{y}$? (NO ! However, will $\beta^* = (\mathbf{X'X})^{-1}\mathbf{X'y}$))

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}(\mathbf{AB})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$$

2 Basic operations in linear algebra

Definition 2.1 Matrix Equality

Matrices $\mathbf{A}$ and $\mathbf{B}$ are **equal** if and only if : (1) they have the same dimension, and, (2) each element $a_{i,k} \in \mathbf{A}$ is equal to the element $b_{i,k} \in \mathbf{B}$ for all i, k. That is,

$$\mathbf{A} = \mathbf{B} \iff a_{i,k} = b_{i,k} \;\; \forall (i \in n, k \in K)$$

Example 2.1 Matrix Equality

If

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 2 & 4 \end{bmatrix}; \;\; \mathbf{B} = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 2 & 4 \end{bmatrix}$$

then, $\mathbf{A} = \mathbf{B}$.

Also, if

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}; \;\; \mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \;\; \textit{and} \;\; \mathbf{X} = \mathbf{A}$$

then, $\{x_{11} = 1, \;\; x_{12} = 0, \;\; x_{21} = -1, \;\; x_{22} = 2\}$.

Definition 2.2 Matrix Addition

Matrix addition is the binary operation that results from adding the element $a_{i,k} \in \mathbf{A}$ to the element $b_{i,k} \in \mathbf{B}$ to obtain a new matrix $\mathbf{C}$ such that the element $c_{i,k} = a_{i,k} + b_{i,k}$. Only matrices with the same dimension are conformable for addition. **Matrix subtraction** is the homologous operation for subtraction.

Example 2.2 Matrix Addition

If $\mathbf{A_{nK}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK} \end{bmatrix}$ and $\mathbf{B_{nK}} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nK} \end{bmatrix}$, then:

$$A_{nK} + B_{nK} = C_{nK} = \begin{bmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \cdots & \alpha_{1K}+\beta_{1K} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \cdots & \alpha_{2K}+\beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}+\beta_{n1} & \alpha_{n2}+\beta_{n2} & \cdots & \alpha_{nK}+\beta_{nK} \end{bmatrix}$$

Also if, $\mathbf{A_{22}} = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$ and $\mathbf{B_{22}} = \begin{bmatrix} 5 & 22 \\ 3 & -1 \end{bmatrix}$, then:

$$A_{22} + B_{22} = C_{22} = \begin{bmatrix} 1+5=6 & -2+22=20 \\ -1+3=2 & 3-1=2 \end{bmatrix} = \begin{bmatrix} 6 & 20 \\ 2 & 2 \end{bmatrix}$$

Finally, note that $\mathbf{A_{22}}+\mathbf{B_{23}}$ is not defined. That's because the matrices are *not conformable* for addition.

Proposition 2.1 Adding the null matrix

The **null matrix** is the additive identity, i.e.,

$$\mathbf{A} + \mathbf{0} = \mathbf{A}$$

. Let $\mathbf{A_{nK}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK} \end{bmatrix}$ and $\mathbf{O_n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$, then:

$$\mathbf{A_{nK}} + \mathbf{0_n} \begin{bmatrix} \alpha_{11}+0 & \alpha_{12}+0 & \cdots & \alpha_{1K}+0 \\ \alpha_{21}+0 & \alpha_{22}+0 & \cdots & \alpha_{2K}+0 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}+0 & \alpha_{n2}+0 & \cdots & \alpha_{nK}+0 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK} \end{bmatrix} = \mathbf{A_{nK}}$$

The addition of matrices is both **commutative** and **associative**,

$$\begin{aligned}\mathbf{A}+\mathbf{B} &= \mathbf{B}+\mathbf{A} && \text{(Commutative Law)}\\ (\mathbf{A}+\mathbf{B})+\mathbf{C} &= \mathbf{A}+(\mathbf{B}+\mathbf{C}) && \text{(Associative Law)}\end{aligned}$$

Proposition 2.2 Matrix Addition: Commutative Law

Let $\mathbf{A_{nK}} = \begin{bmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K}\\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K}\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK}\end{bmatrix}$ and $\mathbf{B_{nK}} = \begin{bmatrix}\beta_{11} & \beta_{12} & \cdots & \beta_{1K}\\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K}\\ \vdots & \vdots & \ddots & \vdots\\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nK}\end{bmatrix}$, then:

$$A_{nK} + B_{nK} = \begin{bmatrix}\alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \cdots & \alpha_{1K}+\beta_{1K}\\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \cdots & \alpha_{2K}+\beta_{2K}\\ \vdots & \vdots & \ddots & \vdots\\ \alpha_{n1}+\beta_{n1} & \alpha_{n2}+\beta_{n2} & \cdots & \alpha_{nK}+\beta_{nK}\end{bmatrix} = C1_{nK}$$

and,

$$B_{nK} + A_{nK} = \begin{bmatrix}\beta_{11}+\alpha_{11} & \beta_{12}+\alpha_{12} & \cdots & \beta_{1K}+\alpha_{1K}\\ \beta_{21}+\alpha_{21} & \beta_{22}+\alpha_{22} & \cdots & \beta_{2K}+\alpha_{2K}\\ \vdots & \vdots & \ddots & \vdots\\ \beta_{n1}+\alpha_{n1} & \beta_{n2}+\alpha_{n2} & \cdots & \beta_{nK}+\alpha_{nK}\end{bmatrix} = C2_{nK}$$

because every $ikth$ element in $C1_{nK}$ is the same in $C2_{nK}$ the two matrices are the same, therefore $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$

Proposition 2.3 Matrix Addition: Associative Law

Let $\mathbf{A_{nK}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK} \end{bmatrix}$, $\mathbf{B_{nK}} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nK} \end{bmatrix}$, and

$\mathbf{C_{nK}} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1K} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nK} \end{bmatrix}$,

then: $(\mathbf{A_{nK}} + \mathbf{B_{nK}}) + \mathbf{C_{nK}} =$

$$\left(\begin{bmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \cdots & \alpha_{1K}+\beta_{1K} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \cdots & \alpha_{2K}+\beta_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}+\beta_{n1} & \alpha_{n2}+\beta_{n2} & \cdots & \alpha_{nK}+\beta_{nK} \end{bmatrix}\right) + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1K} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nK} \end{bmatrix} =$$

$$\begin{bmatrix} \alpha_{11}+\beta_{11}+\gamma_{11} & \alpha_{12}+\beta_{12}+\gamma_{12} & \cdots & \alpha_{1K}+\beta_{1K}+\gamma_{1K} \\ \alpha_{21}+\beta_{21}+\gamma_{21} & \alpha_{22}+\beta_{22}+\gamma_{22} & \cdots & \alpha_{2K}+\beta_{2K}+\gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}+\beta_{n1}+\gamma_{n1} & \alpha_{n2}+\beta_{n2}+\gamma_{n2} & \cdots & \alpha_{nK}+\beta_{nK}+\gamma_{nK} \end{bmatrix}$$

also: $\mathbf{A_{nK}} + (\mathbf{B_{nK}} + \mathbf{C_{nK}}) =$

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1K} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nK} \end{bmatrix} + \left(\begin{bmatrix} \beta_{11}+\gamma_{11} & \beta_{12}+\gamma_{12} & \cdots & \beta_{1K}+\gamma_{1K} \\ \beta_{21}+\gamma_{21} & \beta_{22}+\gamma_{22} & \cdots & \beta_{2K}+\gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1}+\gamma_{n1} & \beta_{n2}+\gamma_{n2} & \cdots & \beta_{nK}+\gamma_{nK} \end{bmatrix}\right) =$$

$$\begin{bmatrix} \alpha_{11}+\beta_{11}+\gamma_{11} & \alpha_{12}+\beta_{12}+\gamma_{12} & \cdots & \alpha_{1K}+\beta_{1K}+\gamma_{1K} \\ \alpha_{21}+\beta_{21}+\gamma_{21} & \alpha_{22}+\beta_{22}+\gamma_{22} & \cdots & \alpha_{2K}+\beta_{2K}+\gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1}+\beta_{n1}+\gamma_{n1} & \alpha_{n2}+\beta_{n2}+\gamma_{n2} & \cdots & \alpha_{nK}+\beta_{nK}+\gamma_{nK} \end{bmatrix}$$

Because the two resulting matrices are equal it has to be true that $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$

Definition 2.3 Transposition

The **transpose** $\mathbf{A}'$ of matrix $\mathbf{A}$ is a matrix whose kth row is the kth column of $\mathbf{A}$.[a]

Formally,

$$\mathbf{B} = \mathbf{A}' \iff b_{i,k} = a_{k,i} \ \ \forall (i \in n, k \in K)$$

[a]Some authors will denote the transpose of $\mathbf{A}$ as $\mathbf{A^T}$ instead of $\mathbf{A}'$.

Example 2.3 Transposition

Let $\mathbf{A_{22}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then $\mathbf{A}' = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$

Let $\mathbf{B_{23}} = \begin{bmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$, then $\mathbf{B}' = \begin{bmatrix} 6 & 3 \\ 5 & 2 \\ 4 & 1 \end{bmatrix}$

Let $\mathbf{C_n} = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix}$, then $\mathbf{C}' = \begin{bmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{bmatrix}$

- By definition every row of $\mathbf{A}'$ is a column of $\mathbf{A}$, then if the dimension of $\mathbf{A}$ is $n \times K$ then the dimension of $\mathbf{A}'$ is $K \times n$.
- Again, by definition of transpose $(\mathbf{A}')' = \mathbf{A}$.
- The transpose of a column vector $\mathbf{a}$ is a row vector $\mathbf{b}$:

Example 2.4 Transpose of vectors

$$\text{If } \ \mathbf{b} = \mathbf{a}' \quad \text{and} \quad \mathbf{a_{n1}} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}, \quad \text{then} \quad \mathbf{b_{1n}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix}$$

- Using the definition of a symmetric matrix we know that,

$$\mathbf{A} \ \text{ is symmetric } \iff \mathbf{A} = \mathbf{A}'$$

Because diagonal matrices are symmetric, the last statement will be true to all diagonal matrices (including the scalar, identity and null matrices).

- The transpose of a lower (upper) triangular matrix is an upper (lower) triangular matrix.
- The transpose is **distributive over the addition of vectors/matrices**.

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$

$$(\mathbf{a} + \mathbf{b})' = \mathbf{a}' + \mathbf{b}'$$

Definition 2.4 Inner product or dot product

The **inner product**, or **dot product**, is a binary operation that results from adding the products of the *ith* elements of two vectors. For two vectors to be conformable for inner product they need to have the same number of elements. Given two vectors $\mathbf{a}$ and $\mathbf{b}$ with the same length (n) their inner product is:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n = \sum_{i=1}^{n} a_i b_i$$

Example 2.5 Dot product

Let $\mathbf{a} = \begin{bmatrix} 10 & 2 & 4 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -1 & 0 & 3 \end{bmatrix}$

Then,

$$\mathbf{a} \cdot \mathbf{b} = 10 \times (-1) + 2 \times (0) + 5 \times 3 = 5$$

Also, let $\mathbf{c} = \begin{bmatrix} -2 & 2 \end{bmatrix}$

Then neither $\mathbf{a} \cdot \mathbf{c}$ nor $\mathbf{b} \cdot \mathbf{c}$ is defined. Because the length of the two vectors do not coincide. If that's the case we say that the vectors are *not conformable* for dot product.

- The result of an inner product is an scalar.
- The dot product is just defined for vectors of the same size, not matrices.
- The inner product is a **commutative** operation

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$

- Is **distributive over vector addition**:

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$$

- The associative property is meaningless for a dot product because the dot product of two vectors is a scalar.

Definition 2.5 Norm

The **Norm** of a vector $\mathbf{z}$, $\|\mathbf{z}\|$, is the (Euclidean) distance between the origin $(0, 0)$ and $\mathbf{z}$. The norm is the square root of the inner product of $\mathbf{z}$ with itself:

$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$

Then,

$d(\mathbf{z}, 0) = \sqrt{(z_1)^2 + (z_2)^2 + \cdots + (z_n)^2} = \sqrt{\sum_{i=1}^{n} z_i^2} = \sqrt{\mathbf{z} \cdot \mathbf{z}} = \|\mathbf{z}\|$

Example 2.6 Norm

Let $\mathbf{z} = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$, then $\|z\| = \sqrt{\mathbf{z} \cdot \mathbf{z}} = \sqrt{1 \times 1 + 0 \times 0 + (-1) \times (-1)} = \sqrt{2}$

- The norm is non-negative.

Definition 2.6 Orthogonality

Two vectors $\mathbf{a}$ and $\mathbf{b}$ are **orthogonal** if they are perpendicular. We denote orthogonality using the $\perp$ symbol. If $\mathbf{a}$ and $\mathbf{b}$ are orthogonal then $\mathbf{a} \perp \mathbf{b}$. The inner product of two orthogonal vectors is zero.

Proposition 2.4 The inner product of orthogonal vectors is zero

Two inner product of orthogonal vectors is $\mathbf{a} \cdot \mathbf{b} = 0$.

Recall the Law of Cosines: *"If a triangle has sides A, B, and C and the angle θ is opposite to the side C*, then

$$c^2 = a^2 + b^2 - 2ab\cos(\theta)$$

where the lowercase letters represent the length of the sides of the triangle.

Let $\mathbf{a}$ be the vector representation of A, $\mathbf{b}$ of B, and $\mathbf{c}$ of C. Then, it has to be true that: $\mathbf{c} = \mathbf{a} - \mathbf{b}$. Then,

$$\|\mathbf{a} - \mathbf{b}\|^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\mathbf{a} \cdot \mathbf{b}$$

Therefore, because $c^2 = \|\mathbf{c}\|^2$ it has to be the case that:

$$\begin{aligned} -2ab\cos(\theta) &= -2\mathbf{a} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{b} &= ab\cos(\theta) \end{aligned}$$

Knowing that $\cos(\theta) = 0$ implies an angle of $\theta = 90°$ between A and B (perpendicular sides), it is clear that if $\mathbf{a} \perp \mathbf{b} \iff \mathbf{a} \cdot \mathbf{b} = 0$

Definition 2.7 Scalar multiplication

A **scalar multiplication** of a matrix (or vector) $\mathbf{A}$ by a scalar $\mathbf{c}$ is a matrix (or vector) $\mathbf{cA}$ of the same dimension of $\mathbf{A}$ whose elements are the product of the scalar and the elements of the original matrix (or vector). That is,

$$\text{If } \mathbf{c} \text{ is a scalar and } \mathbf{A_{nK}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nK} \end{bmatrix} \quad \text{then,}$$

$$\mathbf{cA_{nK}} = \begin{bmatrix} c \times a_{11} & c \times a_{12} & \cdots & c \times a_{1K} \\ c \times a_{21} & c \times a_{22} & \cdots & c \times a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ c \times a_{n1} & c \times a_{n2} & \cdots & c \times a_{nK} \end{bmatrix}$$

Example 2.7 Scalar multiplication

Let $\mathbf{A_{nk}} = \begin{bmatrix} 10 & 2 & -1 \\ 2 & 4 & 1 \end{bmatrix}$, then: If $\mathbf{c} = 2$, $\mathbf{cA_{nK}} = \begin{bmatrix} 20 & 4 & -2 \\ 4 & 8 & 2 \end{bmatrix}$

- If $\mathbf{c} = 0$ then multiplying $\mathbf{c}$ to a square matrix $\mathbf{A}$ will result in a null matrix $\mathbf{0}$ with the dimensions of $\mathbf{A}$.
- If $\mathbf{c} = 1$ then multiplying $\mathbf{c}$ to a matrix $\mathbf{A}$ will result in the same matrix $\mathbf{A}$.
- The scalar multiplication is **commutative**:

$$\mathbf{cA} = \mathbf{Ac}$$

- Is **distributive over vector/matrix addition**, let $\mathbf{A}$ and $\mathbf{B}$ be two matrices of the same size and $\mathbf{a}$ and $\mathbf{b}$ two row (column) vectors of the same length:

$$\mathbf{c(A + B) = cA + cB}$$

$$\mathbf{c(a + b) = ca + cb}$$

Definition 2.8 Matrix multiplication

A **matrix multiplication** of matrices $\mathbf{A_{n\times K}}$ and $\mathbf{B_{K\times m}}$ is a matrix $\mathbf{C_{n\times m}}$ such that the *ikth* element of **C** is the inner product of the *ith* row vector of **A** and the *kth* column vector of **B**. Thus, $\mathbf{c_{i,k}} = \mathbf{a_i} \cdot \mathbf{b_k}$. Two matrices **A** and **B** are **conformable for matrix multiplication AB** if the number of columns in **A** (**premultiplication matrix**) is equal to the number of rows in **B** (**postmultiplication matrix**). This is because every element of the product of **A** and **B** is an inner product, and in order to compute the inner product of two vectors they need to be of the same length. Therefore,

$$\mathbf{A_{n\times K}} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,K} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,K} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$\text{And, } \mathbf{B_{K\times m}} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{K,1} & b_{K,2} & \cdots & b_{K,m} \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & \cdots & b_m \end{bmatrix}$$

$$\underset{(\mathbf{n\times m})}{\mathbf{AB}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_m \end{bmatrix} = \begin{bmatrix} a_1\cdot b_1 & a_1\cdot b_2 & \cdots & a_1\cdot b_m \\ a_2\cdot b_1 & a_2\cdot b_2 & \cdots & a_2\cdot b_m \\ \vdots & \vdots & \ddots & \vdots \\ a_n\cdot b_1 & a_n\cdot b_2 & \cdots & a_n\cdot b_m \end{bmatrix}$$

Example 2.8 Matrix Multiplication

Let $\mathbf{A_{23}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$, and $\mathbf{B_{32}} = \begin{bmatrix} 10 & 2 \\ 1 & 2 \\ 0 & -1 \end{bmatrix}$, then,

$$\mathbf{AB_{22}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 10 & 2 \\ 1 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1\times 10 + 0\times 1 + (-1)\times 0 & 1\times 2 + 0\times 2 + (-1)\times(-1) \\ 0\times 10 + 1\times 1 + 2\times 0 & 0\times 2 + 1\times 2 + 2\times(-1) \end{bmatrix}$$

$$\mathbf{AB_{22}} = \begin{bmatrix} 10 & 3 \\ 1 & 0 \end{bmatrix}$$

- The matrix multiplication is **not commutative**:

$$\mathbf{AB} = \mathbf{C} \not\Rightarrow \mathbf{BA} = \mathbf{C}$$

- Vectors and matrices can be multiplied as long as they satisfy the conformability condition.

- The matrix multiplication is **distributive over vector/matrix addition**:

$$\mathbf{A}(\mathbf{b} + \mathbf{c}) = \mathbf{Ab} + \mathbf{Ac}$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$$

- The matrix multiplication is **associative**:

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

- The matrix multiplication is **distributive over scalar multiplication**:

$$\begin{aligned} \mathbf{c}(\mathbf{AB}) &= (\mathbf{cA})\mathbf{B} \\ &= \mathbf{A}(\mathbf{cB}) \\ &= (\mathbf{AB})\mathbf{c} \\ &= \mathbf{AcB} \end{aligned}$$

- The identity matrix is the **multiplicative identity**. That is, multiplying a matrix $\mathbf{A}$ by the identity $\mathbf{I}$ results in the matrix $\mathbf{A}$, provided that $\mathbf{A}$ and $\mathbf{I}$ are conformable.

- Multiplying by the null matrix results in the null matrix (provided that the matrices are conformable).

- The transpose of a product $(\mathbf{AB})'$ is $\mathbf{B}'\mathbf{A}'$.

- The transpose of an extended product is $(\mathbf{ABC})'$ is $\mathbf{C}'\mathbf{B}'\mathbf{A}'$.

- A **linear combination** is an expression constructed from a set of terms by multiplying each term by a constant and adding the result. Thus, the rows of $\mathbf{C} = \mathbf{Ab}$, where $\mathbf{A_{n\times K}}$ is a matrix and $\mathbf{b_{K\times 1}}$ a vector, are the linear combinations of the columns of $\mathbf{A}$ where the coefficients of the linear combination are the elements of $\mathbf{b}$. Moreover, when computing $\mathbf{C} = \mathbf{AB}$, where

$\mathbf{A_{n\times K}}$ is a matrix and $\mathbf{B_{K\times m}}$ is also a matrix, the columns of $\mathbf{C}$ are a linear combination of the columns of $\mathbf{A}$ where the coefficients are the columns of $\mathbf{B}$.

- Let $\mathbf{a_{n\times 1}}$ and $\mathbf{b_{k\times 1}}$. In order to multiply $\mathbf{a}$ and $\mathbf{b}$ we need to verify that the vectors are conformable. Therefore, we need to transpose $\mathbf{b}$ to obtain $\mathbf{c_{n\times K}} = \mathbf{ab'}$. Also, you can easily verify that $\mathbf{ab'} = \mathbf{ba'}$. If $\mathbf{n} = \mathbf{K}$ then the multiplications $\mathbf{b'a} = \mathbf{a'b}$ are well defined and are equal to the inner (dot) product $\mathbf{a}\cdot\mathbf{b}$.

3 Linear systems of equations

3.1 Representation in matrix form, determinants, adjoints and inverses

A system of n linear equations with K variables can be represented in matrix form as follows:

$$\mathbf{Ax} = \mathbf{b}$$

Where $\mathbf{A}$ is the matrix of coefficients of the system, $\mathbf{x}$ is a vector of variables and $\mathbf{b}$ is a vector of constants. Note that $\mathbf{A}$ and $\mathbf{x}$ need to be conformable for matrix multiplication, i.e. the number of columns in $\mathbf{A}$ has to be equal to the number of rows of $\mathbf{x}$. Finally, the dimension of $\mathbf{b}$ should be the number of rows in A and the number of columns in x (recall this result from matrix multiplication).

Example 3.1 Representation of a linear system in matrix form

Consider the following system of linear equations:

$$\begin{array}{cccccc} x_1 & -2x_2 & +3x_3 & = & 2 \\ -x_1 & & +\frac{1}{2}x_3 & = & 0 \\ & 4x_2 & -x_3 & = & 6 \end{array}$$

We can write the system in matrix form as follows:

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 0 & \frac{1}{2} \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix}$$

Note that by defining,

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 \\ -1 & 0 & \frac{1}{2} \\ 0 & 4 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix}$$

we can write the system as,

$$\mathbf{Ax} = \mathbf{b}$$

- For $\mathbf{Ax} = \mathbf{b}$, if $\mathbf{b} = 0$ we say that the system of equations is **homogeneous**. If $\mathbf{b}$ is nonzero then the system of equations is **nonhomogeneous**
- A **homogeneous** system always has the trivial solution $\mathbf{x} = \mathbf{0}$.

Definition 3.1 Inverse of a Matrix

For a square matrix $\mathbf{A}$, the **inverse**, denoted by $\mathbf{A^{-1}}$, is a matrix such that if $\mathbf{A}$ is premultiplied or postmultiplied by $\mathbf{A^{-1}}$ the result is the identity matrix. That is,

$$\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$$

- The inverse is **only defined for square matrices**.
- A matrix $\mathbf{A}$ is **nonsingular** if and only if its inverse exist. A matrix without an inverse is a **singular** matrix.
- If $\mathbf{A}$ is symmetric then its inverse is also symmetric.
- For square matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{AB})^{-1} = \mathbf{B^{-1}A^{-1}}$, provided that both inverses exist.
- For square matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}(\mathbf{AB})^{-1} = \mathbf{C^{-1}B^{-1}A^{-1}}$ provided that all the inverses exist.

Definition 3.2 Minor and Cofactor of a Matrix

Let $\mathbf{A}$ be a $n \times n$ matrix. Then,

- The (i, j) **minor**, denoted $M_{i,j}$, is the $(n-1) \times (n-1)$ matrix obtained from $\mathbf{A}$ by deleting the i^{th}row and the j^{th} column.
- The (i, j) **cofactor**, denoted $C_{i,j}$, is defined in terms of the minor by $C_{i,j} = (-1)^{i+j} M_{i,j}$

Definition 3.3 Determinant

Let $\mathbf{A}$ be an $n \times n$ matrix with entries a_{ij}.

The **determinant** of a matrix $\mathbf{A_{nn}}$, $|\mathbf{A}|$ (or $\det \mathbf{A}$), is defined as:

For $n = 1$,

$$|\mathbf{A}| = a_{1,1}$$

In general, for $n \geq 2$,

For any row $i = 1, 2, ..., n$,

$$|\mathbf{A}| = a_{i,1}C_{i,1} + a_{i,2}C_{i,2} + ... + a_{i,n}C_{i,n} = \sum_{j=1}^{n} a_{i,j}C_{i,j}$$

where $C_{i,j}$ is the **cofactor** i, j of the matrix $\mathbf{A}$.

This method is called the *cofactor expansion along the i th row.*

Recall, a cofactor $C_{i,j}$ is the determinant of a submatrix of $\mathbf{A}$ obtained by eliminating row i and column j, or **minor** $M_{i,j}$, times $(-1)^{i+j}$.

Thus, determinant of $\mathbf{A}$ in terms of co-factor expansion along 1st row

$$|\mathbf{A}| = a_{1,1}C_{1,1} + a_{1,2}C_{1,2} + ... + a_{1,n}C_{1,n} = \sum_{j=1}^{n} a_{1,j}C_{1,j}$$

can be re-written in terms of minors as follows -

$$|\mathbf{A}| = a_{1,1}M_{1,1} - a_{1,2}M_{1,2} + ...(-1^{1+n})a_{1,n}M_{1,n} = \sum_{j=1}^{n} (-1)^{1+j} a_{1,j}M_{1,j}$$

Alternatively, determinant of $\mathbf{A}$ can be found by cofactor expansion along the j th column for any $j = 1, 2, ..., n$,

$$|\mathbf{A}| = a_{1,j}C_{1,j} + a_{2,j}C_{2,j} + ... + a_{n,j}C_{n,j} = \sum_{i=1}^{n} a_{i,j}C_{i,j}$$

Example 3.2 Determinants

Let $\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 0 & 1 \end{bmatrix}$

To find the determinant of A we need to:

- Find the cofactors (in this case C_{11} and C_{12}, because $n = 2$).
  - To find C_{11} we need to compute the determinant of the submatrix obtained by eliminating row 1 and column 1. Therefore, $C_{11} = |1| = 1$.
  - To find C_{11} we need to compute the determinant of the submatrix obtained by eliminating row 1 and column 2. Therefore, $C_{12} = |0| = 0$.
- Multiply each cofactor by it's corresponding entry from the original matrix and taking into account $(-1)^{i+1}$.

$$|\mathbf{A}| = 2 \times 1 - 4 \times 0 = 2$$

The determinant is a function $det : squarematrices \longrightarrow R$ satisfying the following properties:

1) Doing a row replacement on $\mathbf{A}$ does not change $|\mathbf{A}|$
2) Scaling a row of $\mathbf{A}$ by a scalar c multiplies the determinant by c
3) Swapping two rows of a matrix multiplies the determinant by -1.
4) The determinant of the identity matrix $\mathbf{I_n}$ is equal to 1.

- For a matrix $\mathbf{A}$ of dimension $K \times K$ and scalar c. $|\mathbf{cA}| = \mathbf{c^K}|\mathbf{A}|$
- $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$
- $|\mathbf{A}| = |\mathbf{A}'|$

Definition 3.4 Adjoint matrix

The **adjoint** of matrix $\mathbf{A}$ is a matrix created by multiplying the j, i cofactors of $\mathbf{A}$.

$$\text{adj } \mathbf{A} = C_{j,i}$$

Example 3.3 Adjoint matrix

Let $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$

The cofactors are: $M_{11} = 4$ $M_{21} = 3$, $M_{12} = 1$, and $M_{22} = 2$.

Then,

$$\text{adj } \mathbf{A} = \begin{bmatrix} (-1)^{1+1}M_{11} & (-1)^{2+1}M_{21} \\ (-1)^{1+2}M_{12} & (-1)^{2+2}M_{22} \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix}$$

- The inverse of a square matrix $\mathbf{A}$ is given by $\mathbf{A^{-1}} = \frac{1}{\det \mathbf{A}}\text{adj } \mathbf{A}$. Therefore, if $\det \mathbf{A} = 0$ the inverse of $\mathbf{A}$ does not exist and $\mathbf{A}$ is singular.

Example 3.4 Finding the inverse using the adjoint matrix and determinant

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

In the previous example we found that:

$$\text{adj } \mathbf{A} = \begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix}$$

The determinant of $\mathbf{A}$ is:

$$\det \mathbf{A} = 2 \times 4 - 1 \times 3 = 5$$

Therefore,

$$\mathbf{A^{-1}} = \frac{1}{\det \mathbf{A}}\text{adj } \mathbf{A} = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.6 \\ -0.2 & 0.4 \end{bmatrix}$$

Note that $\mathbf{AA^{-1}} = I$.

$$\mathbf{AA^{-1}} = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 0.8 & -0.6 \\ -0.2 & 0.4 \end{bmatrix} = \begin{bmatrix} 2 \times 0.8 - 3 \times 0.2 & 2 \times (-0.6) + 3 \times 0.4 \\ 1 \times 0.8 - 4 \times 0.2 & 1 \times (-0.6) + 4 \times 0.4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Definition 3.5 Solution of a Nonhomogeneous Linear System of Equations

Given the **nonhomogenous system of linear equations**,

$$\mathbf{Ax} = \mathbf{b}$$

If $\mathbf{A}$ is nonsingular then,

$$\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$$

Is a **solution** to the given system.

The previous result can be easily verified by premultiplying the original system of equations by the inverse of the coefficient matrix.

$$\begin{aligned}
\mathbf{Ax} &= \mathbf{b} \\
\mathbf{A}^{-1}\mathbf{Ax} &= \mathbf{A}^{-1}\mathbf{b} \\
\mathbf{Ix} &= \mathbf{A}^{-1}\mathbf{b} \\
\mathbf{x}^* &= \mathbf{A}^{-1}\mathbf{b}
\end{aligned}$$

To guarantee that $\mathbf{A}^{-1}$ exist we need to assume that $\mathbf{A}$ is nonsingular.

Example 3.5 Solution of a linear system

Consider the system from Example (3.1). The inverse of $\mathbf{A}$ is,

$$\mathbf{A^{-1}} = \begin{bmatrix} 1/6 & -5/6 & 1/12 \\ 1/12 & 1/12 & 7/24 \\ 1/3 & 1/3 & 1/6 \end{bmatrix}$$

Note that,

$$\mathbf{A^{-1}A} = \begin{bmatrix} 1/6 & -5/6 & 1/12 \\ 1/12 & 1/12 & 7/24 \\ 1/3 & 1/3 & 1/6 \end{bmatrix} \begin{bmatrix} 1 & -2 & 3 \\ -1 & 0 & \frac{1}{2} \\ 0 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}$$

and,

$$\mathbf{x}^* = \mathbf{A^{-1}b} = \begin{bmatrix} 1/6 & -5/6 & 1/12 \\ 1/12 & 1/12 & 7/24 \\ 1/3 & 1/3 & 1/6 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} 20 \\ 1 \\ -16 \end{bmatrix}$$

You can verify that $\mathbf{x}^*$ is the actual solution of the system by multiplying $\mathbf{Ax}^*$ and checking that is equal to $\mathbf{b}$.

3.2 Linear independence and solutions to linear systems

Definition 3.6 Linear transformation (Linear Map)

A **linear trasnformation** (linear map, linear function) is a function from the vectors space into the vectors space that **preserves the operations of addition and scalar multiplication**. Therefore, for a linear transformation L, vectors $\mathbf{a}$ and $\mathbf{b}$, and scalar α it has to be true that:

$$\begin{aligned} L(\mathbf{a}+\mathbf{b}) &= L(\mathbf{a}) + L(\mathbf{b}) \quad \text{(additivity)} \\ L(\alpha\mathbf{a}) &= \alpha L(\mathbf{a}) \quad \text{(scalar multiplication)} \end{aligned}$$

Matrix multiplications are linear transformations.[a]

[a] The second property (scalar) is also known as *homogeneity of degree* 1.

Example 3.6 Matrices and linear transformation

Let $\mathbf{A_{23}} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 3 & 3 \end{bmatrix}$; and $\mathbf{x_{31}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Note that the premultiplication by $\mathbf{A}$ is a linear transformation from R^3 to R^2 (it takes a vector with $n = 3$ and returns one with $n = 2$). $\mathbf{Ax} = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 & 14 \end{bmatrix}$

Definition 3.7 Linear combination

A **linear combination** of vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector $\mathbf{c}$ that is obtained by sum of the product of each original vector by a scalar.

$$\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$$

In general the linear combination $\mathbf{y}$ of vectors $\{\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_K}\}$ by scalars $\{\lambda_1, \lambda_2, ...\lambda_K\}$ is:

$$\mathbf{y} = \sum_{i=1}^{K} \lambda_i \mathbf{x_i}$$

Definition 3.8 Linear Dependence

A set of $k \geq 2$ vectors is **linearly dependent** if at least one of the vectors in the set is a linear combination of the others.

Example 3.7 Linearly dependent vectors

Consider vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$,

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

then, because

$$\mathbf{a} + \mathbf{b} = \mathbf{c}$$

we say that $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are linearly dependent. Any matrix made out of this three vectors will be linearly dependent, the order of the vectors in the matrix does not matter. Recall that a matrix is just a set of vectors.

Definition 3.9 Linear Independence

A set of vectors **A** is **linearly independent**, if and only if, given a a row of coefficients α, the only solution to:

$$\begin{aligned}\alpha\mathbf{A} &= 0\\ \alpha_1\mathbf{a_1}+\alpha_2\mathbf{a_2}+...+\alpha_K\mathbf{a_K} &= 0\end{aligned}$$

is,

$$\alpha_1=\alpha_2=...=\alpha_K=0$$

Systems of linear equations may have:

1. **A unique solution**
2. **No solution**
3. **Infinite solutions**

A solution to a linear system $\mathbf{Ax}=\mathbf{b}$ is a *point* x^* such that the K equations are simultaneously satisfied; from a geometric point of view, for the case of $K=2$, this represents the intersection of two lines, for $K=3$ the intersection of three planes, for $K>3$ the intersection of K hyperplanes.

Consider the case of $K=2$; if we have two equations and two variables, we know that:

- If the slope of the two equations **are different** then there is **only one point in which the two lines intersects.**
- If the slopes **are the same** we may have **no intersection** (parallel lines) or **infinite intersection** (the two equations describe the same line).

For example,

$$\begin{aligned}a_{1,1}x+a_{1,2}y &= b_1\\ a_{2,1}x+a_{2,2}y &= b_2\end{aligned}$$

which can be rewritten as,

$$\begin{aligned}y &= \frac{b_1}{a_{1,2}}-\frac{a_{1,1}}{a_{1,2}}x\\ y &= \frac{b_2}{a_{2,1}}-\frac{a_{2,1}}{a_{2,2}}x\end{aligned}$$

If,

$$\frac{a_{1,1}}{a_{1,2}} = \frac{a_{2,1}}{a_{2,2}} \rightarrow a_{1,1}a_{2,2} - a_{1,2}a_{2,1} = 0$$

Then the system may have no or infinite solutions. Therefore, if $a_{1,1}a_{2,2} - a_{1,2}a_{2,1} \neq 0$ **we can guarantee that the system has a solution**. That is, by comparing the slopes of two linear equations we can *determine* if the system has a unique solution or not.

Recall that the analogous concept to the slope of a curve is the gradient (see lecture notes on static optimization for a definition of a gradient). The same logic that we use for $K = 2$ can be applied to $K > 2$ to discuss the existence of an unique solution to the system, if at least the gradients of two of the K hyperplanes are identical then the system do not have an unique solution. This is the basis of the idea of **determinants**.

- For $K = 2$, note that $|\mathbf{A}| \neq 0 \leftrightarrow \frac{a_{1,1}}{a_{1,2}} \neq \frac{a_{2,1}}{a_{2,2}}$. That is, if the determinant of a matrix is non-zero we know that a system has an unique solution. You can also verify that this is true for $K > 2$.
- A matrix with a linearly dependent column or row (vector) will have a determinant equal to zero; this imply that the matrix is singular, the inverse and, therefore, the solution is not determined.

Example 3.8 The determinant of a 2×2 linearly dependent matrix

Consider vectors $\mathbf{x} = \begin{bmatrix} 0.5 & 0.2 \end{bmatrix}$ and $\mathbf{y} = 2\mathbf{x} = \begin{bmatrix} 1 & 0.4 \end{bmatrix}$

Now consider the matrix:

$$\mathbf{A} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.2 \\ 1 & 0.4 \end{bmatrix}$$

The determinant is given by:

$$\det \mathbf{A} = 0.5 \times 0.4 - 1 \times 0.2 = 0$$

In general for $n = 2$, Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$ and $\mathbf{y} = \lambda \mathbf{x} = \begin{bmatrix} \lambda x_1 & \lambda x_2 \end{bmatrix}$. Where λ is a scalar.

$$\mathbf{A} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ \lambda x_1 & \lambda x_2 \end{bmatrix}$$

The determinant is:

$$\det \mathbf{A} = \lambda x_1 x_2 - \lambda x_1 x_2 = 0$$

You can prove that this is true also for $n > 2$. It will take a while to do it, but it's possible.

4 Eigenvalues, diagonalization, rank and trace

Definition 4.1 Eigenvalues

The **eigenvalues** of the square matrix $\mathbf{A}$ are defined as the set of λ that satisfy:

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{c} = 0$$

Where $\mathbf{c}$ is a column vector.

- The eigenvalues of a matrix can be negative/zero/positive and real/complex.
- The system $(\mathbf{A} - \lambda \mathbf{I})\mathbf{c} = 0$ has a solution only if the determinant $|\mathbf{A} - \lambda \mathbf{I}| = 0$. This determinant is known as the **characteristic equation** of $\mathbf{A}$.
- A $K \times K$ matrix has K eigenvalues.

- An **eigenvector** is a vector $\mathbf{c}$ that results from solving $(\mathbf{A}-\lambda\mathbf{I})\mathbf{c} = 0$. A square matrix $\mathbf{A}$ with K columns has at least K eigenvectors. The eigenvectors are also known as **characteristic vectors**.

- If $\mathbf{c}$ is a null vector we say that it is a **trivial** solution, because it will satisfy $(\mathbf{A} - \lambda\mathbf{I})\mathbf{c} = 0$ for any set of λ.

Example 4.1 Finding eigenvalues

Consider the matrix,

$$\mathbf{A} = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

then,

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 5-\lambda & 1 \\ 4 & 2-\lambda \end{bmatrix}$$

Computing the determinant we get the following characteristic equation,

$$|\mathbf{A} - \lambda\mathbf{I}| = (5-\lambda)(2-\lambda) - (4)(1) = \lambda^2 - 7\lambda + 6$$

By making the determinant equal to zero, the roots of the resulting second-degree polynomial are $\lambda_1 = 6$ and $\lambda_2 = 1$. This are the eigenvalues of $\mathbf{A}$. Using each eigenvalue we can obtain the correspondent eigenvector of $\mathbf{A}$.

for $\lambda_1 = 6$

$$(\mathbf{A} - 6\mathbf{I})\mathbf{c} = \begin{bmatrix} -1 & 1 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{array}{rcrl} -c_1 & + & c_2 & = 0 \\ 4c_1 & - & 4c_2 & = 0 \end{array}$$

Therefore, the set of vectors such that $c_1 = c_2$ are the set of eigenvectors for $\lambda_1 = 6$.

for $\lambda_2 = 1$

$$(\mathbf{A} - 1\mathbf{I})\mathbf{c} = \begin{bmatrix} 4 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{array}{rcrl} 4c_1 & + & c_2 & = 0 \\ 4c_1 & + & c_2 & = 0 \end{array}$$

The the set of solutions to the previous system is given by $(c_1, -4c_1)$. This is the set of eigenvectors for $\lambda_2 = 1$.

Proposition 4.1 Different eigenvalues corresponds to different eigenvectors

Proof by contradiction.

Consider the matrix $\mathbf{A}$, the set of eigenvalues $\{\lambda_1, \lambda_2, ...\lambda_n\}$ and the corresponding set of vectors $\{\mathbf{c}_1, \mathbf{c}_2, ...\mathbf{c}_n\}$. Assume that:

$$\lambda_i \neq \lambda_j \quad \forall\, i, j \in N$$

and,

$$\exists\, c_i = c_j \quad \text{for some } \ i, j \in N$$

Then,

If $c_i = c_j$ it follows that $\mathbf{A}c_i = \mathbf{A}c_j$. Using the definition of eigenvalues $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{c_i} = 0 \rightarrow \mathbf{A}\mathbf{c_i} = \lambda_i \mathbf{c_i}$. Which implies that:

$$\begin{aligned} \mathbf{A}c_i &= \mathbf{A}c_j \\ \lambda_i c_i &= \lambda_j c_j \\ \lambda_i c_i &= \lambda_j c_i \qquad (c_i = c_j) \\ (\lambda_i - \lambda_j)c_i &= 0 \end{aligned}$$

which contradicts the initial assumption that $\lambda_i \neq \lambda_j$. ■

Proposition 4.2 Different eigenvectors are orthogonal (for symmetric matrices)

Direct proof Recall that if two vectors are different then they have different eigenvalues, so we can safely assume that if eigenvectors $c_i \neq c_j$ the corresponding eigenvalues have to be $\lambda_i \neq \lambda_j$. Then by definition

$$\begin{aligned}(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{c_i} &= 0 \\ (\mathbf{A} - \lambda_j \mathbf{I})\mathbf{c_j} &= 0\end{aligned}$$

Multiplying both expressions by $\mathbf{c_j'}$ and $\mathbf{c_i'}$ respectively.

$$\begin{aligned}\mathbf{c_j'}(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{c_i} &= 0 \\ \mathbf{c_i'}(\mathbf{A} - \lambda_j \mathbf{I})\mathbf{c_j} &= 0\end{aligned}$$

Substracting the two equations yields,

$$\begin{aligned}\mathbf{c_j'}(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{c_i} - \mathbf{c_i'}(\mathbf{A} - \lambda_j \mathbf{I})\mathbf{c_j} &= 0 \\ \mathbf{c_i'}(\mathbf{A} - \lambda_i \mathbf{I} - \mathbf{A} + \lambda_j \mathbf{I})\mathbf{c_j} &= 0 \quad \text{(By symmetry)} \\ (\lambda_j - \lambda_i)\mathbf{c_i'}\mathbf{c_j} &= 0\end{aligned}$$

Because we start assuming $\lambda_j \neq \lambda_i$ it has to be the case that $\mathbf{c_i'}\mathbf{c_j} = 0$. Recall that $\mathbf{c_i'}\mathbf{c_j} = c_i \cdot c_j = 0 \rightarrow c_i \perp c_j$. ■

Definition 4.2 Diagonalization of a Matrix

The **diagonalization** of matrix **A** is obtained by,

$$\mathbf{C^{-1}AC} = \mathbf{C^{-1}C\Lambda} = \mathbf{\Lambda}$$

where $\mathbf{C}$ is a $K \times K$ matrix such the i column vector $\mathbf{c_i} \in \mathbf{C}$ is the eigenvector that corresponds to the eigenvalue i of $\mathbf{A}$. And $\mathbf{\Lambda}$ is a diagonal matrix with K and the main diagonal is equal to the vector of eigenvalues.

- The first step to prove that $\mathbf{C^{-1}AC} = \mathbf{C^{-1}C\Lambda}$, is to realize that $\mathbf{AC} = \mathbf{C\Lambda}$ and premultiply by $\mathbf{C^{-1}}$. Note that $\mathbf{AC} = \mathbf{C\Lambda}$ is just the representation of the system $\mathbf{Ac_i} = \lambda_i \mathbf{c_i}$ which is what we are doing when finding the eigenvectors. Therefore, $\mathbf{AC} = \mathbf{C\Lambda}$ has to be true.

- Then, because $\mathbf{C^{-1}C} = \mathbf{I}$, $\mathbf{C'C\Lambda} = \mathbf{\Lambda}$.

- Representing $\mathbf{A}$ as $\mathbf{C\Lambda C^{-1}}$ is called **spectral decomposition**.

- Becauase $\mathbf{C^{-1}AC} = \mathbf{\Lambda}$ implies $|\mathbf{C^{-1}AC}| = |\mathbf{\Lambda}|$ and $|\mathbf{C^{-1}AC}| = |\mathbf{C^{-1}}||\mathbf{A}||\mathbf{C}| = |\mathbf{C^{-1}}||\mathbf{C}||\mathbf{A}| = |\mathbf{C^{-1}C}||\mathbf{A}| = |\mathbf{I}||\mathbf{A}| = \mathbf{1}|\mathbf{A}| = |\mathbf{A}|$. Then it has to be the case that $|\mathbf{A}| = |\mathbf{\Lambda}|$. That is, the determinant of a matrix is equal to the determinant of its diagonalization. Also, **the determinant of a matrix is equal to the product of its eigenvalues.**

- If a matrix has **any eigenvalue equal to zero the determinant of the matrix is equal to zero.** And, consequently, the matrix is singular and not invertible.

- If **all eigenvalues are positive the determinant** of the matrix **is positive.**

- If **the number of negative eigenvalues is even the determinant is positive**, if the is number of negative eigenvalues **is odd** then **the determinant is negative**.

Definition 4.3 Trace of a matrix

The **trace** of a matrix $\mathbf{A}$ is the sum of its diagonal elements,

$$\mathbf{tr}(\mathbf{A}) = \sum_{i=1}^{K} a_{i,i}$$

- For a matrix $\mathbf{A}$ and scalar c, $\mathbf{tr}(c\mathbf{A}) = c\mathbf{tr}(\mathbf{A})$.

- $\mathbf{tr}(\mathbf{A}) = \mathbf{tr}(\mathbf{A'})$

- $\mathbf{tr}(\mathbf{A} + \mathbf{B}) = \mathbf{tr}(\mathbf{A}) + \mathbf{tr}(\mathbf{B})$

- $\mathbf{tr}(\mathbf{I_K}) = K$

- $\mathbf{tr}(\mathbf{AB}) = \mathbf{tr}(\mathbf{BA})$

- $\mathbf{tr}(\mathbf{A'A}) = \sum_{i=1}^{K} a_{i,i}a_{i,i} = \sum_{i=1}^{K} a_{i,i}^2$

- $\mathbf{tr}(\mathbf{C^{-1}AC}) = \mathbf{tr}(\mathbf{C^{-1}C\Lambda}) = \mathbf{tr}(\mathbf{\Lambda})$, and; because $\mathbf{tr}(\mathbf{C^{-1}AC}) = \mathbf{tr}(\mathbf{C^{-1}CA}) = \mathbf{tr}(\mathbf{AC^{-1}C}) = \mathbf{tr}(\mathbf{A})$, we know that $\mathbf{tr}(\mathbf{A}) = \mathbf{tr}(\mathbf{\Lambda})$. That is, the trace of a square matrix is equal to the trace of its diagonalization. From this it follows that **the trace of a matrix is equal to the sum of its eigenvalues**.

Definition 4.4 Column and Row Rank

The **column rank** of $\mathbf{A}$ is the largest number of linearly independent columns of $\mathbf{A}$.The **row rank** of $\mathbf{A}$ is the largest number of linearly independent rows of $\mathbf{A}$.

Theorem 4.1 Equality of Column and Row Rank

The **column rank** and **column row** of a matrix are equal.

- A matrix $\mathbf{A}$ is **full column (row) rank** if the number of columns (rows) is equal to the column (row) rank.
- Because the column rank and row rank are equal we won't make a distinction among them from now on and simply use the term **rank** of a matrix.
- $\mathbf{rank}(\mathbf{A}_{n\times K}) = \mathbf{rank}(\mathbf{A}'_{K\times n}) \leq \min(n, K)$.
- $\mathbf{rank}(\mathbf{AB}) \leq \min(\mathbf{rank}(\mathbf{A}), \mathbf{rank}(\mathbf{B}))$.
- $\mathbf{rank}(\mathbf{A}) = \mathbf{rank}(\mathbf{A}'\mathbf{A}) = \mathbf{rank}(\mathbf{AA}')$.
- For a $K \times K$ full rank matrix $\mathbf{C}$, $\mathbf{rank}(\mathbf{C}^{-1}\mathbf{AC}) = \mathbf{rank}(\mathbf{A})$. And because $\mathbf{rank}(\mathbf{C}'\mathbf{AC}) = \mathbf{rank}(\mathbf{\Lambda})$ it has to be the case that $\mathbf{rank}(\mathbf{A}) = \mathbf{rank}(\mathbf{\Lambda})$. Therefore, if any eigenvalue of matrix $\mathbf{A}$ is zero the matrix won't have full rank (some columns are linearly dependent). On the other hand, if all eigenvalues are different than zero the matrix has full rank. If a matrix is full rank its **determinant is not zero**.

5 Quadratic forms and definite matrices

Definition 5.1 Quadratic form

The **quadratic form** q is defined as:

$$q = \mathbf{x}'\mathbf{A}\mathbf{x}$$

where $\mathbf{x}$ is a $K \times 1$ nonzero vector and $\mathbf{A}$ is a $K \times K$ symmetric matrix. Note that we can also represent q as a sum.

$$q = \sum_{i=1}^{K}\sum_{j=1}^{K} x_i x_j A_{i,j}$$

- If $q = \mathbf{x}'\mathbf{A}\mathbf{x} > (<)0$ for all nonzero $\mathbf{x}$, then $\mathbf{A}$ is **positive (negative) definite**.
- If $q = \mathbf{x}'\mathbf{A}\mathbf{x} \geq (\leq)0$ for all nonzero $\mathbf{x}$, then $\mathbf{A}$ is **positive (negative) semidefinite**.

- For the $K \times K$ symmetric matrix $\mathbf{A}$ the spectral decomposition is $\mathbf{C\Lambda C^{-1}}$. Therefore,

$$\begin{aligned}
\mathbf{x'Ax} &= \mathbf{x'C\Lambda C^{-1}x} \\
&= \mathbf{y'\Lambda y} \\
&= \sum_{i=1}^{K} \lambda_i y_i^2
\end{aligned}$$

Theorem 5.1 Definiteness and Eigenvalues

Let $\mathbf{A}$ be a $K \times K$ symmetric matrix. If all eigenvalues are positive (negative) then the matrix $\mathbf{A}$ is **positive (negative) definite**. If some roots are zero, and the rest are all positive (negative) then the matrix $\mathbf{A}$ is **positive (negative) semidefinite**. If $\mathbf{A}$ has both positive and negative eigenvalues then $\mathbf{A}$ is **indefinite**.

- If $\mathbf{A}$ positive (semi)definite, then $|\mathbf{A}|(\geq) > 0$.
- If $\mathbf{\Lambda}$ is positive definite, then $\mathbf{A}$ is also positive definite.
- The identity matrix $\mathbf{I}$ is **positive definite**.

6 Calculus and liner algebra

The following results of calculus and linear algebra are useful to remember (specially for econometrics).

- For $\mathbf{y} = \mathbf{Ax}$. Because, $\dfrac{\partial y_i}{\partial \mathbf{x}} = \dfrac{\partial a_i'\mathbf{x}}{\partial x} = a_i'$. Therefore, $\dfrac{\partial \mathbf{y}}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{Ax}}{\partial \mathbf{x}} = \mathbf{A}'$.
- For $\mathbf{y} = \mathbf{x'Ax}$, $\dfrac{\partial \mathbf{y}}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{x'Ax}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}')\mathbf{x}$.
- For the quadratic form $\mathbf{y} = \mathbf{x'Ax}$, $\dfrac{\partial \mathbf{y}}{\partial \mathbf{x}} = \dfrac{\partial \mathbf{x'Ax}}{\partial \mathbf{x}} = 2\mathbf{Ax}$.
- For the quadratic form $\mathbf{y} = \mathbf{x'Ax}$, $\dfrac{\partial \mathbf{y}}{\partial \mathbf{A}} = \dfrac{\partial \mathbf{x'Ax}}{\partial \mathbf{A}} = \mathbf{xx}'$.