Linear Algebra
Edps/Soc 584 and Psych 594
Applied Multivariate Statistics

Carolyn J. Anderson

Department of Educational Psychology

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Outline

- Motivation
- Vectors
- Matrices
- Random Vectors & Matrices
- Useful Matrix Formulas
- Matrix Software (SAS/IML, MATLAB, & R)

Reading & Resources:

- Johnson & Wichern, Chapter 1, Chapter 2, supplement
- Chapter by Larry Hubert linked to course web-site
- Matrix reference manual at http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html
- Matrix cookbook at http://matrixcookbook.com

Motivation

Data Array:

$$\underbrace{n \times p}_{\text{“size” of matrix } \boldsymbol{X}} \quad \text{matrix} \quad \boldsymbol{X} \qquad\qquad (\text{cases} \times \text{variables})$$

Why learn about linear algebra (& geometry)?

- Provides a useful mathematical form.
- Simplifies: shorthand with which extensive material can be condenses into a compact form.
- Perform matrix operations in multivariate analysis.
- Needed for more advanced statistics and multivariate methods

So we need to study matrices & operations and properties of solutions.

Vectors

Vectors: An array $\boldsymbol{x}$ of n real numbers is called a vector & written as

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{or} \quad \boldsymbol{x}' = (x_1, x_2, \ldots, x_n)$$

- Vectors are denoted by lower case bold letters (i.e., when hand written, put a $\sim$ underneath the letter).
- Typically, vectors are columns
- Transpose of a (column) vector is $\boldsymbol{x}'$ is a row vector.
- A vector consists of n scalars (numbers).
- A vector is a matrix with one column, $(n \times 1)$.

- Outline

Motivation

Vectors
- Vectors
- Vector as a Point
- Definition of Equal Vectors
- Multiplication of Vector by a Scalar
- Reflection
- Vector Addition
- Geometry of Vector Addition
- Vector Subtraction
- Length of Vector
- Generalization to n-Dimensional Space
- Length $\longrightarrow$ Distance
- Distance between Two Arbitrary Points
- Squared Distances
- Direction: Angles between two vectors
- Trigonometry Review
- Example: Computing angle between vectors
- Inner Product of Two Vectors
- Generalization to n dimensions
- Linear Dependence
- Linear Independence
- Geometry of Linear (In)Dependence
- More about Basis
- Projection of a Vector
- Projection of a Vector

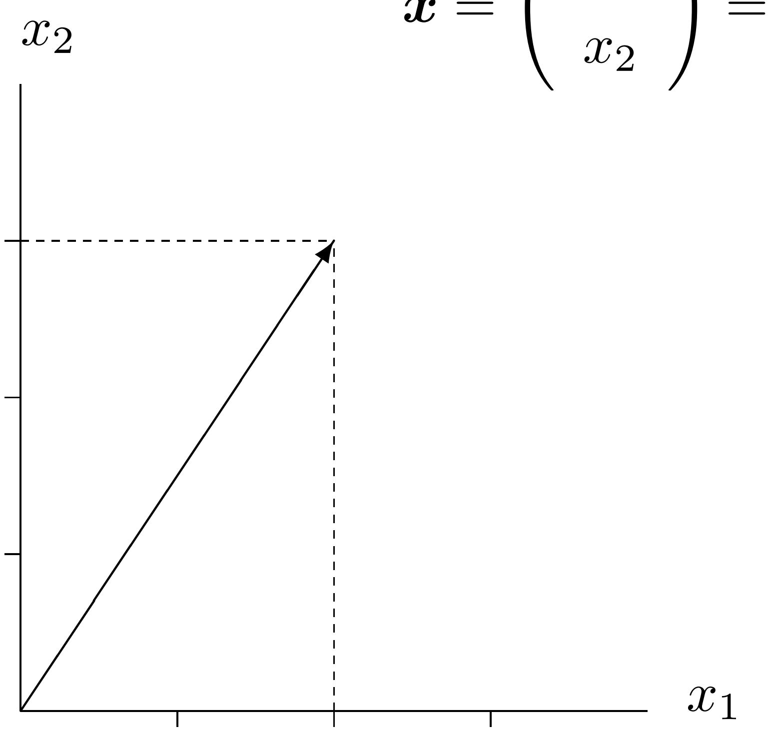
Vector as a Point

A Vector is a point in n-dimensional space represented by an arrow originating at zero.

All vectors have a Length and a Direction.

e.g.,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$



Definition of Equal Vectors

Two vectors (matrices) are equal if they have the same number of rows, the same number of columns, and the corresponding entries in each are equal

$$\boldsymbol{a} = \boldsymbol{b} \qquad \text{means} \qquad a_{ij} = b_{ij} \qquad \text{for all } i \ \& \ j$$

OR

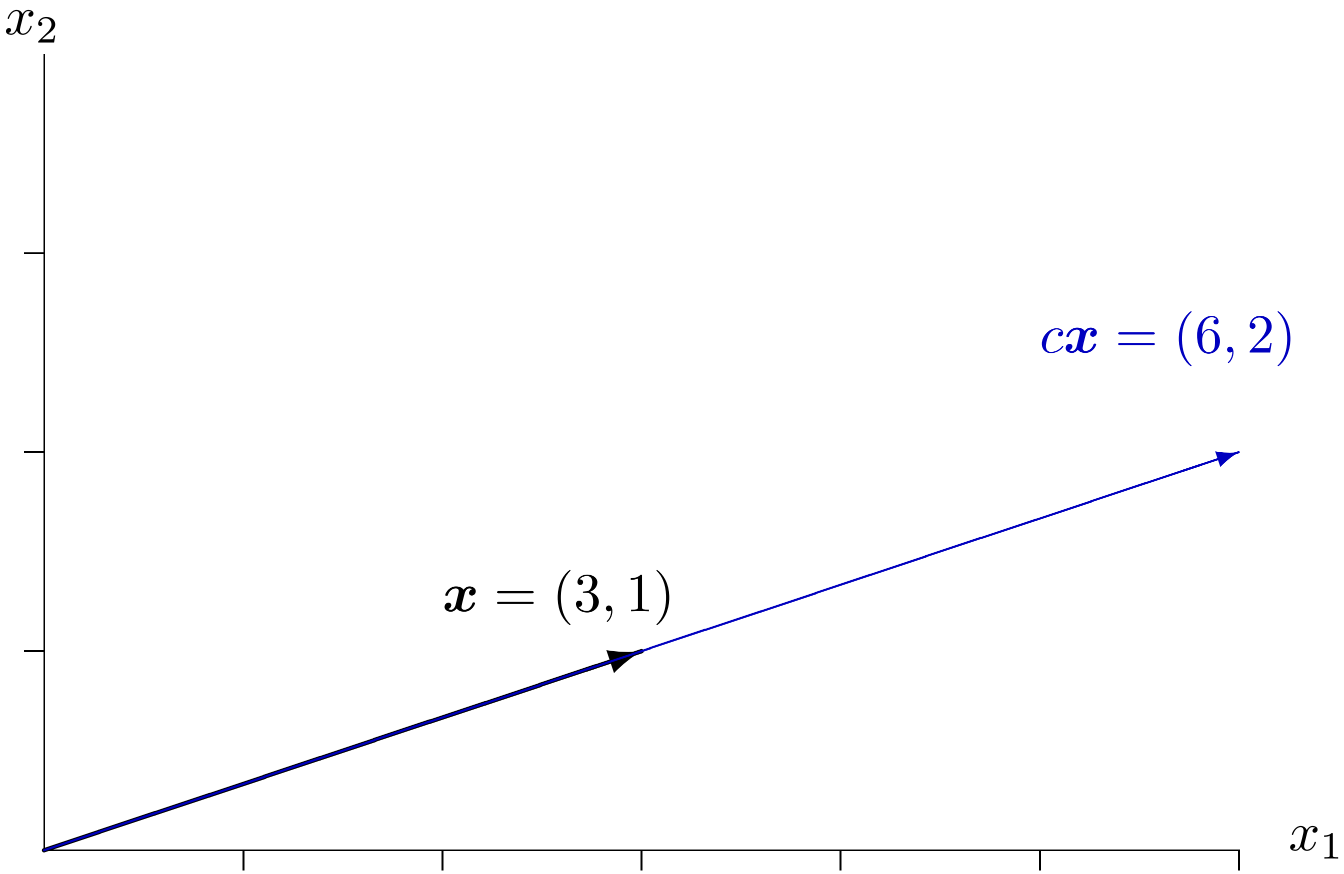
Two vectors are equal (identical) if they both have the same length and direction.

Multiplication of Vector by a Scalar

Suppose that c is some number (a scalar). For scalar multiplication, multiply each entry of $\boldsymbol{x}$ by c:

$$c\boldsymbol{x}' = (cx_1, cx_2, \ldots, cx_n)$$

Example: Let $\boldsymbol{x}' = (3,1)$ and $c = 2$

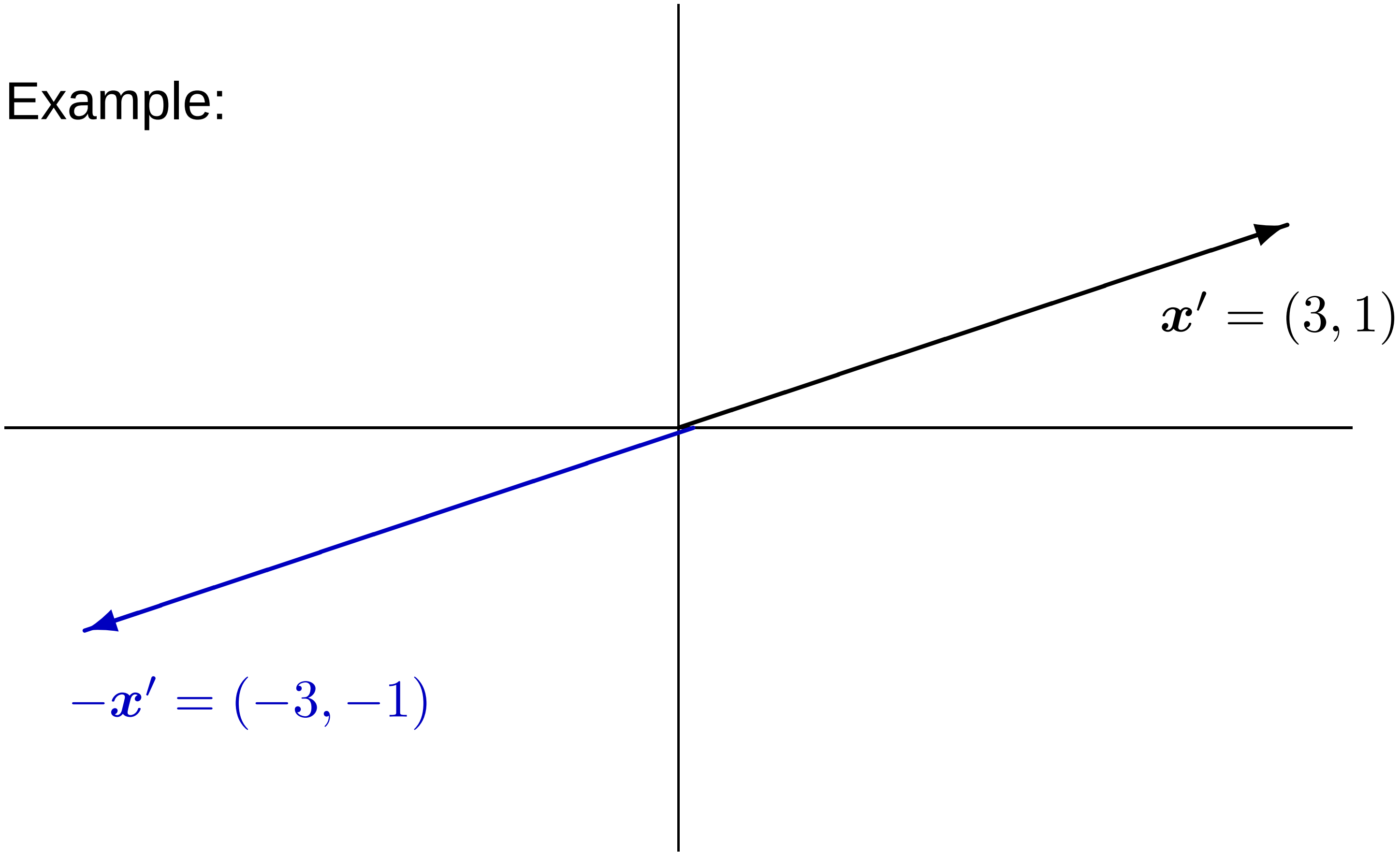


Reflection

Multiplication of a vector by $c = -1$

$$-\boldsymbol{x}' = (-x_1, -x_2, \ldots, -x_n)$$

Example:



Vector Addition

The sum of two vectors, each with the same number of elements (i.e., same size), element by element:

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

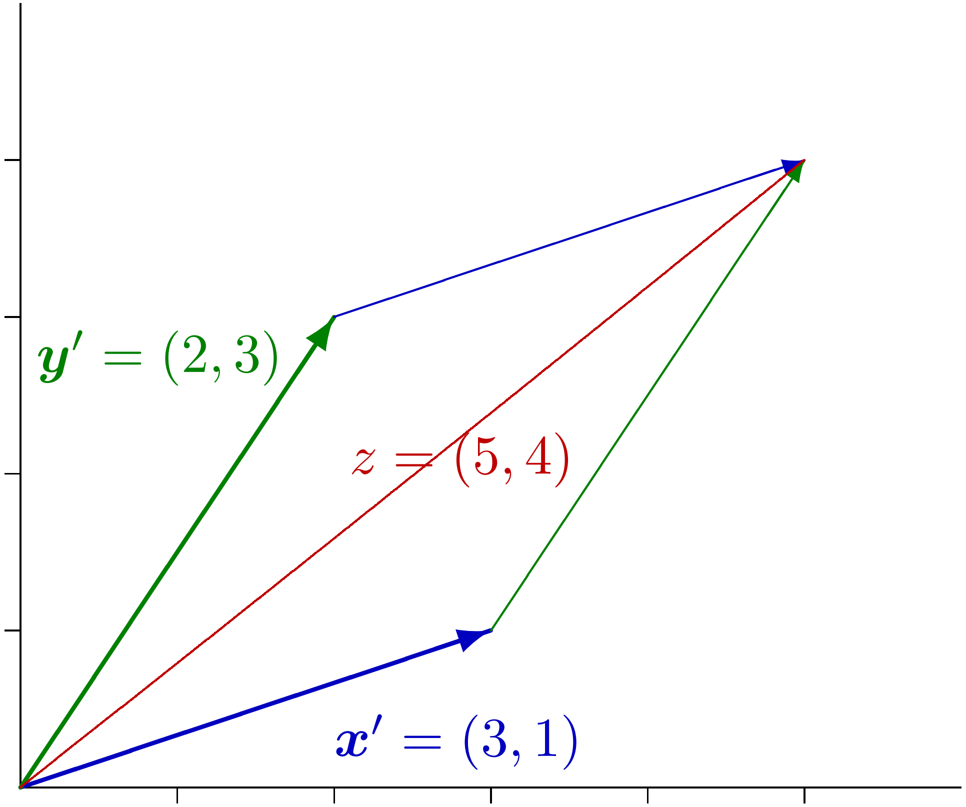
e.g., $\boldsymbol{x}' = (3,1)$ and $\boldsymbol{y}' = (2,3)$

$$\boldsymbol{x} + \boldsymbol{y} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} =$$

Geometry of Vector Addition

Form a parallelogram by sliding the origin of one vector to the end of the other.

The sum is the diagonal of the parallelogram. e.g.,

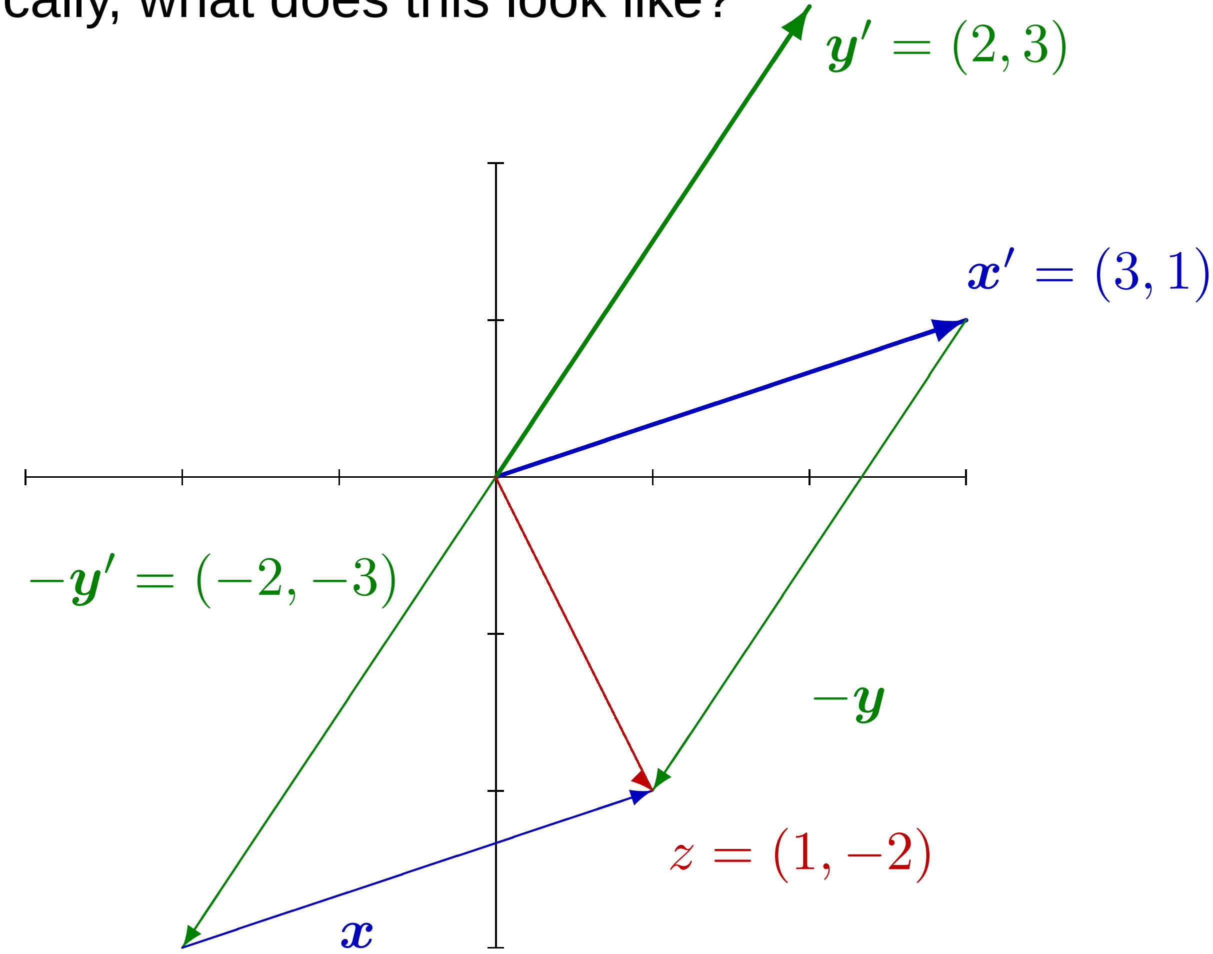


Vector Subtraction

Element by element addition

$$\boldsymbol{x}' - \boldsymbol{y}' = (x_1 - y_1, x_2 - y_2, \ldots, x_I - y_I)$$

Geometrically, what does this look like?

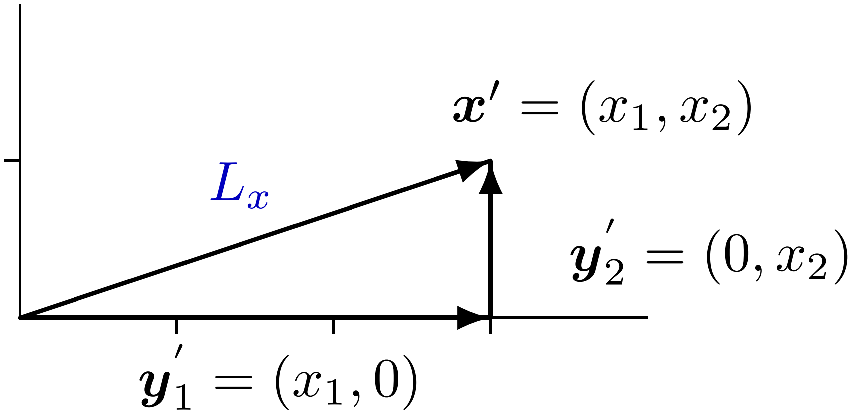


Length of Vector

Remember the Pythagorean Theorem?

Let

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = \boldsymbol{y}_1 + \boldsymbol{y}_2$$



$$L_{\boldsymbol{x}} = \sqrt{x_1^2 + x_2^2}$$

Generalization to n-Dimensional Space

Let

$$\boldsymbol{x}' = (x_1, x_2, \ldots, x_n)$$

And the length of $\boldsymbol{x}$ is

$$L_x = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\sum_{j=1}^{n} x_j^2}$$

Implications for scalar multiplication:

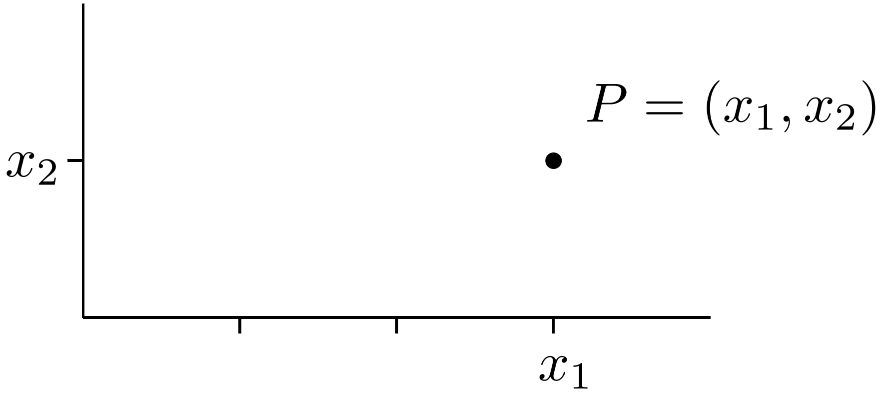
$$\begin{aligned} c\boldsymbol{x} \longrightarrow L_{c\boldsymbol{x}} &= \sqrt{c^2x_1^2 + c^2x_2^2 + \ldots + c^2x_n^2} \\ &= |c|\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} \\ &= |c|L_x \end{aligned}$$

- If $c < 0$, then change direction.
- If $c = -1$, then change direction but not length.

Length $\longrightarrow$ Distance

Distance between points.

Suppose we have a 2-dimensional space and a point $P = (x_1, x_2)$



Distance from the origin $0 = (0, 0)$ to point P is

$$d(0, P) = \sqrt{x_1^2 + x_2^2} = L_x$$

If point P has n coordinates, $P = (x_1, x_2, \ldots, x_n)$ then the "straight-line" distance from the origin, $0 = (0, 0, \ldots, 0)$ to P is

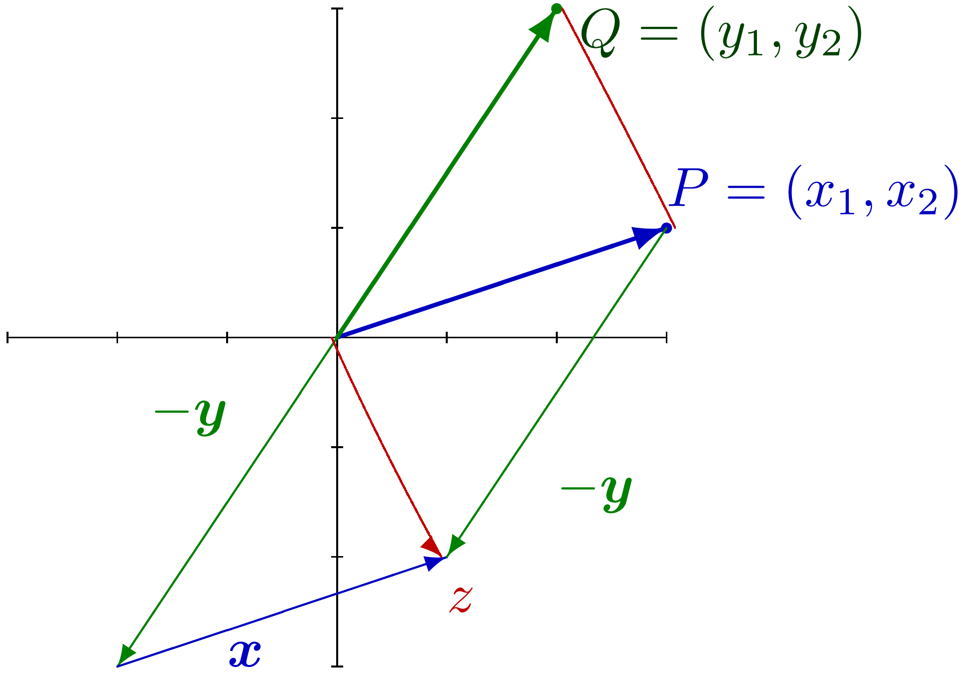
$$d(0, P) = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

Distance between Two Arbitrary Points

The “straight-line” distance between two arbitrary points $P = (x_1, x_2, \ldots, x_n)$ and $Q = (y_1, y_2, \ldots, y_n)$ is

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

Why? Show this for 2-Dimensional case:



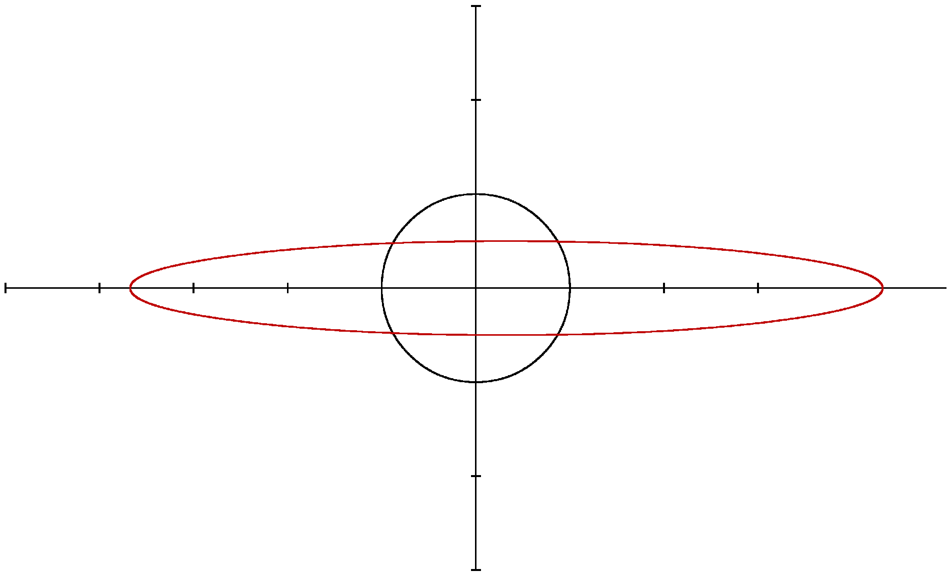
The length of $\boldsymbol{x} - \boldsymbol{y}$ is the distance between $P - Q$.
e.g., $\boldsymbol{x} = (3, 1)$, $\boldsymbol{y} = (2, 3)$, and $\boldsymbol{z}$

Squared Distances

All points that are a constant (squared) distance, say c^2 from the origin $\mathbf{0}$ satisfy the equation

$$d^2(\mathbf{0}, P) = x_1^2 + x_2^2 + \ldots + x_n^2 = c^2$$

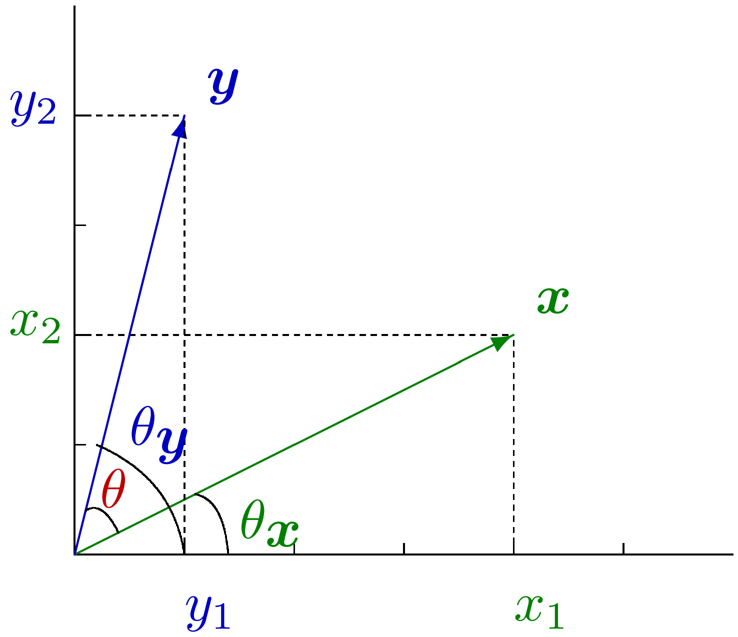
This is the equation for a sphere in n-dimensional space in (un-weighted) Euclidean space.



Statistical Distance: Weight scores by a variable's standard deviation (more about these later). This yields an ellipsoid, which has important statistical connections.

Direction: Angles between two vectors

$\theta =$ the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$



What do we know or can compute?

- x_1, x_2, y_1, and y_2
- $L_{\boldsymbol{x}}$ and $L_{\boldsymbol{y}}$
- $\theta_{\boldsymbol{y}}$ and $\theta_{\boldsymbol{x}}$.
- And $\theta = \theta_{\boldsymbol{y}} - \theta_{\boldsymbol{x}}$.

Trigonometry Review

Remember this?

$$\cos(\theta_{\boldsymbol{x}}) = \frac{x_1}{L_{\boldsymbol{x}}} \qquad \cos(\theta_{\boldsymbol{y}}) = \frac{y_1}{L_{\boldsymbol{y}}}$$

and

$$\sin(\theta_{\boldsymbol{x}}) = \frac{x_2}{L_{\boldsymbol{x}}} \qquad \sin(\theta_{\boldsymbol{y}}) = \frac{y_2}{L_{\boldsymbol{y}}}$$

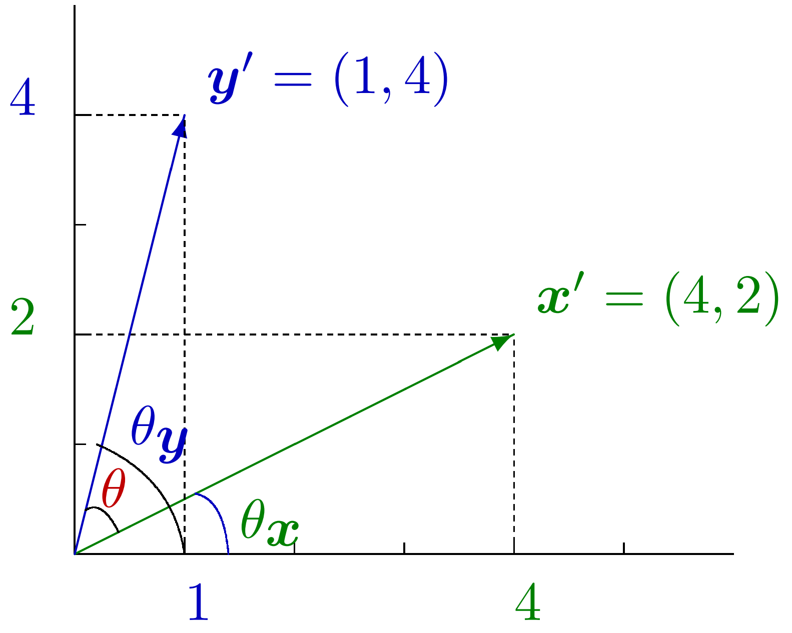
And

$$\begin{aligned} \cos(\theta) &= \cos(\theta_{\boldsymbol{y}} - \theta_{\boldsymbol{x}}) \\ &= \cos(\theta_{\boldsymbol{y}})\cos(\theta_{\boldsymbol{x}}) + \sin(\theta_{\boldsymbol{y}})\sin(\theta_{\boldsymbol{x}}) \end{aligned}$$

So

$$\begin{aligned} \cos(\theta) &= \left(\frac{y_1}{L_{\boldsymbol{y}}}\right)\left(\frac{x_1}{L_{\boldsymbol{x}}}\right) + \left(\frac{y_2}{L_{\boldsymbol{y}}}\right)\left(\frac{x_2}{L_{\boldsymbol{y}}}\right) \\ &= \frac{x_1 y_1 + x_2 y_2}{L_{\boldsymbol{x}} L_{\boldsymbol{y}}} \end{aligned}$$

Example: Computing angle between vectors



$$L_x = \sqrt{4^2 + 2^2} = \sqrt{20} = 4.472136 \qquad L_y = \sqrt{1^2 + 4^2} = \sqrt{17} = 4.12311$$

$$\cos(\theta) = \left(\frac{y_1}{L_y}\right)\left(\frac{x_1}{L_x}\right) + \left(\frac{y_2}{L_y}\right)\left(\frac{x_2}{L_x}\right)$$

$$=$$

Answer: $\cos^{-1}(.65079) = 49.40^\circ$

Inner Product of Two Vectors

Note:

$$\cos(\theta) = \frac{(x_1y_1 + x_2y_2)}{\sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}}$$

The “inner product” of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ equals $(x_1y_1 + x_2y_2)$.

- Also called the “dot product”.
- Denoted by $\boldsymbol{x}'\boldsymbol{y}$ or $\boldsymbol{y}'\boldsymbol{x}$ (order doesn’t matter).
- $\boldsymbol{x}'\boldsymbol{y} = x_1y_1 + x_2y_2$
- $\boldsymbol{x}'\boldsymbol{x} = x_1^2 + x_2^2 = L_{\boldsymbol{x}}^2$
- $\boldsymbol{y}'\boldsymbol{y} = y_1^2 + y_2^2 = L_{\boldsymbol{y}}^2$
- $\cos(\theta) = 0$ when $\theta = 90^\circ \Longleftrightarrow$
  $\boldsymbol{x}$ and $\boldsymbol{y}$ are perpendicular when $\boldsymbol{x}'\boldsymbol{y} = 0$.

Generalization to n dimensions

If $\boldsymbol{x}$ and $\boldsymbol{y}$ are $(n \times 1)$ vectors , i.e.,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Then the inner product between $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = (x_1y_1 + x_2y_2 + \ldots + x_ny_n) = \text{ a scalar}$$

$$\text{Length of } \boldsymbol{x} = L_{\boldsymbol{x}} = \sqrt{\boldsymbol{x}'\boldsymbol{x}}$$

$$\cos(\theta) = \frac{\boldsymbol{x}'\boldsymbol{y}}{L_{\boldsymbol{x}}L_{\boldsymbol{y}}} = \frac{\boldsymbol{x}'\boldsymbol{y}}{\sqrt{\boldsymbol{x}'\boldsymbol{x}}\sqrt{\boldsymbol{y}'\boldsymbol{y}}}$$

$\boldsymbol{x}'\boldsymbol{y} = 0$ if and only if $\boldsymbol{x}$ & $\boldsymbol{y}$ are perpendicular

Linear Dependence

Linearly Dependent: A set of vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, . . ., $\boldsymbol{x}_k$ is said to be linearly dependent if there exists constants c_1, c_2, . . ., c_k not all equal to zero such that

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \ldots + c_k\boldsymbol{x}_k = \boldsymbol{0}$$

If vectors are linearly dependent, then one vector in the set can be written as a function (linear combination) of the others; that is,

$$x_i = -\frac{1}{c_i}\left(\sum_{k\neq i} c_k\boldsymbol{x}_k\right)$$

e.g.,

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad x_3 = \begin{pmatrix} -4 \\ 2 \\ -2 \\ 4 \end{pmatrix}$$

If $c_1 = 1$, $c_2 = -5$ and $c_3 = -1$, then $1\boldsymbol{x}_1 - 5\boldsymbol{x}_2 - 1\boldsymbol{x}_3 = 0$

Linear Independence

If there are no constants $c_1, c_2, \ldots, c_k$ not all equal to zero such that

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \ldots + c_k\boldsymbol{x}_k = \mathbf{0}$$

Then the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k$ are linearly independent.

Vectors are linearly independent if you cannot write one as a linear combination of the others

Example

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

The only constants that make the condition true is

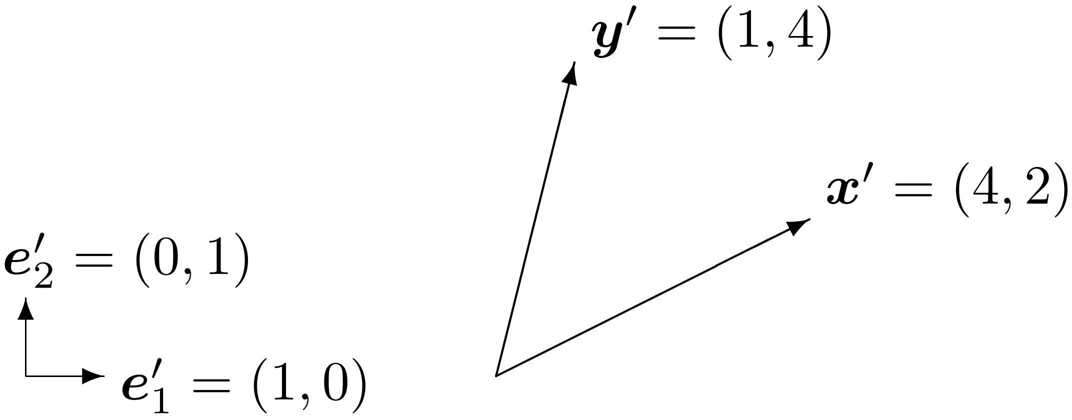
$$c_1 = c_2 = c_k = 0$$

If all vectors in a set are $(k \times 1)$, at most how many vectors could possibly be in a set of linearly independent vectors?

Geometry of Linear (In)Dependence

A basis is a set of linearly independent vectors and a basis defines a space.

- For example, $\boldsymbol{e}_1' = (1,0)$ and $\boldsymbol{e}_2' = (0,1)$ define the space in a plane — these are special ones, a “standard basis”.
- The dimension of the space defined by a set of vectors equals the number of linearly independent vectors in the set.



Any third vector in either of these spaces must be a linear combination of the other two? Try $\boldsymbol{z}' = (6,2)$

More about Basis

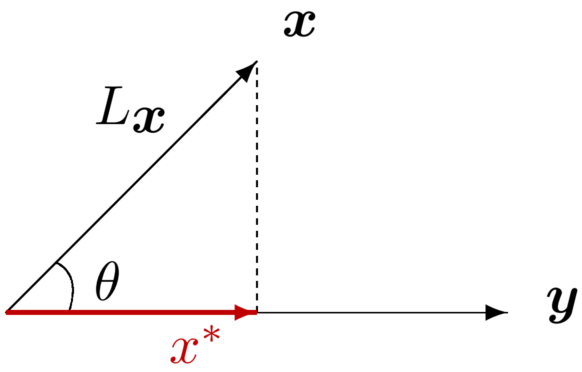
- An example of a standard basis is the set of vectors $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_p$ where $\boldsymbol{e}_k$ equals a $(p \times 1)$ vector with all 0 except for a 1 in the k^{th} position.

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \boldsymbol{e}_p = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

- These vectors are all orthogonal: $\boldsymbol{e}_k'\boldsymbol{e}_l = 0$
- They all have length equal to 1: $L_{\boldsymbol{e}_k} = \sqrt{\boldsymbol{e}_k'\boldsymbol{e}_k} = 1$
- Any basis can always be transformed to an orthogonal one (This can be done using the Grahm-Schmidt orthogonalization process).

Projection of a Vector

Projection of $\boldsymbol{x}$ onto $\boldsymbol{y}$:



The projection of $\boldsymbol{x}$ onto $\boldsymbol{y}$ is in the same direction as $\boldsymbol{y}$, but has a different length.

- Length of projection:
  - We know that $\cos(\theta) = L_{x^*}/L_{\boldsymbol{x}}$.
  - We also know that $\cos(\theta) = (\boldsymbol{x}'\boldsymbol{y})/(L_{\boldsymbol{x}}L_{\boldsymbol{y}})$
  - So the length of the projection is

$$L_{\boldsymbol{x}^*} = L_{\boldsymbol{x}}\cos(\theta) = L_{\boldsymbol{x}}\left(\frac{\boldsymbol{x}'\boldsymbol{y}}{L_{\boldsymbol{x}}L_{\boldsymbol{y}}}\right) = \frac{\boldsymbol{x}'\boldsymbol{y}}{L_{\boldsymbol{y}}}$$

- For the direction, find a unit vector in the direction of $\boldsymbol{y}$,

$$\frac{1}{L_{\boldsymbol{y}}}\boldsymbol{y}$$

● Outline

Motivation

Vectors
● Vectors
● Vector as a Point
● Definition of Equal Vectors
● Multiplication of Vector by a Scalar
● Reflection
● Vector Addition
● Geometry of Vector Addition
● Vector Subtraction
● Length of Vector
● Generalization to n-Dimensional Space
● Length $\longrightarrow$ Distance
● Distance between Two Arbitrary Points
● Squared Distances
● Direction: Angles between two vectors
● Trigonometry Review
● Example: Computing angle between vectors
● Inner Product of Two Vectors
● Generalization to n dimensions
● Linear Dependence
● Linear Independence
● Geometry of Linear (In)Dependence
● More about Basis
● Projection of a Vector
● Projection of a Vector

Projection of a Vector (continued)

- Multiply the length of the projection $L_{\boldsymbol{x}^*}$ (a scalar) times the unit vector in the direction of $\boldsymbol{y}$ (recall what happens when you multiply a vector by a scalar).

- The projection of $\boldsymbol{x}$ on $\boldsymbol{y}$ is

$$\boldsymbol{x}^* = \underbrace{\left(\frac{\boldsymbol{x}'\boldsymbol{y}}{L_{\boldsymbol{y}}}\right)}_{\text{scalar}} \underbrace{\left(\frac{\boldsymbol{y}}{L_{\boldsymbol{y}}}\right)}_{\text{unit vector in same direction as } \boldsymbol{y}}$$

$$= \underbrace{\frac{\boldsymbol{x}'\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{y}}}_{\text{scalar}} \quad \underbrace{\boldsymbol{y}}_{\text{vector}}$$

- One use of this concept: observations (e.g., $\boldsymbol{x}$) can be broken down into the sum of mean (projection) plus deviation or residual (vertical distance from $\boldsymbol{x}$ to end of $\boldsymbol{x}^*$).

Zero and One Vector

● Outline

Motivation

Vectors

● Vectors
● Vector as a Point
● Definition of Equal Vectors
● Multiplication of Vector by a Scalar
● Reflection
● Vector Addition
● Geometry of Vector Addition
● Vector Subtraction
● Length of Vector
● Generalization to n-Dimensional Space
● Length ⟶ Distance
● Distance between Two Arbitrary Points
● Squared Distances
● Direction: Angles between two vectors
● Trigonometry Review
● Example: Computing angle between vectors
● Inner Product of Two Vectors
● Generalization to n dimensions
● Linear Dependence
● Linear Independence
● Geometry of Linear (In)Dependence
● More about Basis
● Projection of a Vector
● Projection of a Vector

The null vector contains all 0's,

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

One use: null hypothesis values when testing multiple means.

Vector of One's

$$\mathbf{1}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Some of the many uses:

- Sum up elements of a vector or sum rows or columns of a matrix.
- Important when we talk about geometry of the mean.

Matrices

- A Matrix is any rectangular array of real numbers. We'll use upper case bold Roman letters to represent them.
- An arbitrary matrix with n rows and p columns

$$\underset{(n \times p)}{\boldsymbol{A}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$$

- The first subscript indicates the row and second indicates a column.
- A matrix with n rows and p columns is said to have order $(n \times p)$.
- Another way to indicate a matrix (and it's order) is

$$\boldsymbol{A} = \{a_{ij}\}_{n \times p}$$

Matrix Operations & Definitions of Special Ones

- Transpose operation is indicated by $'$ (e.g., $\boldsymbol{A}'$).

$$\boldsymbol{A}' = \boldsymbol{B} \Longleftrightarrow b_{ij} = a_{ji}$$

Transpose changes the rows to columns and columns to rows

- Sometimes the letter T is used to indicate transpose, e.g., $\boldsymbol{A}^{\mathtt{T}}$,

Example:

$$\boldsymbol{A}_{(4\times 2)} = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 5 \\ 3 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{A}'_{(2\times 4)} =$$

Equal Matrices

Suppose we have two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same $n \times p$ order:

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}_{n\times p} \qquad \text{and} \qquad \mathbf{B} = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}_{n\times p}$$

As a definition for equality of two matrices of the same order (and for which it only makes sense to talk about equality), we have:

$$\mathbf{A} = \mathbf{B} \quad \textit{if and only if} \quad a_{ij} = b_{ij} \text{ for all } i \text{ and } j.$$

“if and only if” statement (sometimes abbreviated as “iff”) implies two conditions:

1. If $\mathbf{A} = \mathbf{B}$, then $a_{ij} = b_{ij}$ for all i and j, and
2. if $a_{ij} = b_{ij}$ for all i and j, then $\mathbf{A} = \mathbf{B}$.

Any definition by its very nature implies an “if and only if” statement.

Scalar Multiplication

- $c\boldsymbol{A}$ where c is a scalar (constant) and $\boldsymbol{A}$ is an arbitrary matrix of any size.

- Multiply each element of $\boldsymbol{A}$ by the scalar c.

$$c\mathbf{A} = \begin{pmatrix} ca_{11} & \cdots & ca_{1p} \\ \vdots & \ddots & \vdots \\ ca_{n1} & \cdots & ca_{np} \end{pmatrix} = c \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}$$

- $\boldsymbol{B} = c\boldsymbol{A} \Longrightarrow ca_{ij}$.

- Example where $c = 3$:

$$\boldsymbol{B} = 3\boldsymbol{A} = 3 \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 5 \\ 3 & 0 \end{pmatrix}$$

Matrix Addition

- Addition is element-by-element.

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} \Longrightarrow c_{ij} = a_{ij} + b_{ij}$$

- Two matrices must have the same order (referred to as conformal for addition).

- Example:

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 5 \\ 3 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -6 \\ 3 & 1 \\ 3 & -2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -4 \\ 8 & 2 \\ 6 & 3 \\ 4 & 0 \end{pmatrix}$$

- Matrix addition is commutative (i.e., order does not matter): $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$.

- Matrix addition is associative: $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$

Matrix Multiplication

- Let $\boldsymbol{A}_{(n\times k)}$ and $\boldsymbol{B}_{(k\times p)}$.
- For $\boldsymbol{C} = \boldsymbol{AB}$,

$$\begin{aligned} c_{ij} &= \sum_{l=1}^{k} a_{il} b_{lj} \\ &= \boldsymbol{a}_i' \boldsymbol{b}_j \\ &= \text{the inner product between the } i^{th} \text{ row of } \boldsymbol{A} \text{ and the} \\ & \quad\, j^{th} \text{ column of } \boldsymbol{B} \end{aligned}$$

$$= \left[\boxed{a_{i1} \quad a_{i2} \ldots a_{ik}} \right] \left[\boxed{\begin{matrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{kj} \end{matrix}} \right]$$

- The number of columns of $\boldsymbol{A}$ must equal the number of rows of $\boldsymbol{B}$.

Matrix Multiplication

Example: $\boldsymbol{A}_{(2\times 3)}\boldsymbol{B}_{(3\times 2)}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 5 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 4 & 3 \\ -2 & 1 \end{pmatrix} =$$

$$=$$

$$=$$

- In General, matrix multiplication is <u>not commutative</u>; it may not even be possible to calculate both $\boldsymbol{AB}$ and $\boldsymbol{BA}$. e.g.,

$$\boldsymbol{B}_{(3\times 2)}\boldsymbol{A}_{(2\times 3)} = \begin{pmatrix} 2 & 4 & 6 \\ 19 & 11 & 18 \\ 3 & -3 & -4 \end{pmatrix} \neq \boldsymbol{A}_{(2\times 3)}\boldsymbol{B}_{(3\times 2)}$$

Matrix Multiplication is Associative

Matrix multiplication is associative:

$$\boldsymbol{ABC} = (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$$

Using $\boldsymbol{A}$ and $\boldsymbol{B}$ from the previous slide and $\boldsymbol{C}$,

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 1 & 2 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 2 & 0 \\ 4 & 3 \\ -2 & 1 \end{pmatrix}, \qquad \boldsymbol{C} = \begin{pmatrix} 5 & 2 \\ 3 & 10 \end{pmatrix}$$

$$(\boldsymbol{AB})\boldsymbol{C} = \begin{pmatrix} 4 & 9 \\ 10 & 5 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 3 & 10 \end{pmatrix} = \begin{pmatrix} 47 & 98 \\ 65 & 70 \end{pmatrix}$$

$$\boldsymbol{A}(\boldsymbol{BC}) = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 1 & 2 \end{pmatrix} \begin{pmatrix} 10 & 4 \\ 29 & 38 \\ -7 & 6 \end{pmatrix} = \begin{pmatrix} 47 & 98 \\ 65 & 70 \end{pmatrix}$$

More Properties of Matrix Operations

- Matrix multiplication is Distributive
  - Left distributive:

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

  - Right distributive:

$$\boldsymbol{AC} + \boldsymbol{BC} = (\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C}$$

- Properties of transposes:
  - Product of transposes

$$(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}' \qquad \text{and } (\boldsymbol{ABC})' = \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'$$

    etc.
  - Transpose of a transpose equals the un-transposed matrix:

$$(\boldsymbol{A}')' = \boldsymbol{A}$$

  - Transpose of a sum equals the sum of the transposes:

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

Square and Symmetric Matrices

A Square Matrix has the same number of rows as columns. e.g.,

$$\boldsymbol{A}_{(2\times 2)} = \begin{pmatrix} 1 & 2 \\ 5 & 7 \end{pmatrix}$$

A Symmetric Matrix is a square matrix where the rows equal the columns

$$\boldsymbol{A} = \boldsymbol{A}' \qquad \text{i.e., } a_{ij} = a_{ji}$$

Alternatively, the upper triangle equals the transpose of the lower triangle (and visa versa).

e.g.,

$$\boldsymbol{A} = \boldsymbol{A}' = \begin{pmatrix} 9 & \color{blue}{1} & \color{green}{3} \\ \color{blue}{1} & 3 & \color{red}{4} \\ \color{green}{3} & \color{red}{4} & 10 \end{pmatrix}$$

Trace of a Square Matrix

$$\mathbf{A} = \mathbf{A}' = \begin{pmatrix} 1 & 4 & 3 \\ 4 & 7 & -1 \\ 3 & -1 & 3 \end{pmatrix}$$

- The "main" or "principal diagonal" entries are in red.
- The sum of main diagonal entries of a square matrix is called the trace; thus,

$$\text{trace}(\mathbf{A}_{(p\times p)}) \equiv \text{tr}(\mathbf{A}) = a_{11} + \cdots + a_{pp}$$

- Our example,

$$\text{tr}(\mathbf{A}) = 1 + 7 + 3 = 11$$

Null and Diagonal Matrices

A Null Matrix has all elements equal to 0.

$$\mathbf{0}_{(n\times p)} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

A Diagonal Matrix is a square matrix where the only non-zero elements are on the diagonal, e.g.,

$$\boldsymbol{D}_{(p\times p)} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{pp} \end{pmatrix}$$

Sometimes it's written as diag$(a_{11}, a_{22}, \dots, a_{pp})$

Pre-Multiplication by a Diagonal Matrix

When you “pre-multiply” a matrix by a diagonal matrix it has the effect of multiplying each row by the corresponding diagonal, e.g.,

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{pp} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pp} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \dots & a_{11}b_{1p} \\ a_{22}b_{21} & a_{22}b_{22} & \dots & A_{22}b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{pp}b_{p1} & a_{pp}b_{p2} & \dots & a_{pp}b_{pp} \end{pmatrix} = \begin{pmatrix} a_{11}\boldsymbol{b}_1' \\ a_{22}\boldsymbol{b}_2' \\ \vdots \\ a_{pp}\boldsymbol{b}_p' \end{pmatrix}$$

Post-Multiplication by a Diagonal Matrix

When you "post-multiply" a matrix by a diagonal matrix it has the effect of multiplying each column by the corresponding diagonal, e.g.,

$$\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pp} \end{pmatrix} \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{pp} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} & a_{22}b_{12} & \dots & a_{pp}b_{1p} \\ a_{11}b_{21} & a_{22}b_{22} & \dots & a_{pp}b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & a_{22}b_{p2} & \dots & a_{pp}b_{pp} \end{pmatrix} = (a_{11}\boldsymbol{b}_1, a_{22}\boldsymbol{b}_2, \dots, a_{pp}\boldsymbol{b}_p)$$

Identity Matrix

The Identity Matrix is a diagonal matrix with all entries equal to 1, e.g.,

$$\boldsymbol{I}_p = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

$\boldsymbol{I}_p$ is call the identity matrix because

$$\boldsymbol{I}_p \boldsymbol{A}_{(p\times p)} = \boldsymbol{A}_{(p\times p)} \boldsymbol{I}_p = \boldsymbol{A}_{(p\times p)}$$

e.g.,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & 5 & 2 \end{pmatrix}$$

Often see $\sigma^2 \boldsymbol{I}$ or $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$.

Determinant of a Matrix

The determinant of a square matrix, $\boldsymbol{A}_{p\times p}$, is a scalar that is associated with that matrix. Determinants are denoted by $|\boldsymbol{A}|$ or $\det(\boldsymbol{A})$.

How to compute the determinant of

- A 1×1 matrix:

$$|a| = a$$

- A 2×2 matrix:

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow |\boldsymbol{A}| = ad - bc$$

- A 3×3 matrix:

$$\boldsymbol{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \longrightarrow |\boldsymbol{A}| = aei + dhc + gfb - (ceg + fha + idb)$$

Determinant of a $(p \times p)$ Matrix

For $p > 3$, this requires the use a recursive process, but first need some definitions:

- Let $\boldsymbol{A}_{uv}$ is be the $(p-1) \times (p-1)$ submatrix of $\boldsymbol{A}$ constructed by deleting the u^{th} row and v^{th} column of $\mathbf{A}$
- Minor entry of a_{uv} is the $\det(\mathbf{A}_{uv})$
- The signed minor of $(-1)^{u+v}\det(\mathbf{A}_{uv})$ is the cofactor of a_{uv}.

The recursive algorithm:

- (1) Chose some row (or column) either arbitrarily or for convenience.
- (2) Find the cofactors for the entries in it.
- (3) The cofactors would then be weighted by the relevant entries and summed. i.e.,

$$\det(\boldsymbol{A}) = \sum_{u=1}^{p} a_{uv}(-1)^{u+v}\det(\boldsymbol{A}_{uv})$$

Example: Determinant of a (4×4) matrix

$$\boldsymbol{A} = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 2 & 5 & 1 & 0 \\ 1 & 3 & 2 & 1 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

I'll use the 2nd row, which I highlighted in red:

$$\begin{aligned} \det(A) &= 2(-1)^{1+2}\begin{vmatrix} 3 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 4 \end{vmatrix} + 5(-1)^{2+2}\begin{vmatrix} 4 & 2 & 1 \\ 1 & 2 & 1 \\ 3 & 1 & 4 \end{vmatrix} \\ &\quad + 1(-1)^{2+3}\begin{vmatrix} 4 & 3 & 1 \\ 1 & 3 & 1 \\ 3 & 2 & 4 \end{vmatrix} + 0(-1)^{2+4}\begin{vmatrix} 4 & 3 & 2 \\ 1 & 3 & 2 \\ 3 & 2 & 1 \end{vmatrix} \\ &= -2(0) + 5(21) - 1(30) + 0 \\ &= 75 \end{aligned}$$

More About Determinants

- The determinant of a diagonal matrix is equal to the product of the diagonal elements:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{pp} \end{pmatrix} \longrightarrow \det(\boldsymbol{A}) = \prod_{j=1}^{p} a_{jj}$$

- The determinant of a upper– (or lower–) triangular matrix equals the product of the diagonal elements:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ 0 & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{pp} \end{pmatrix} \longrightarrow \det(\boldsymbol{A}) = \prod_{j=1}^{p} a_{jj}$$

Properties of Determinants

- If one row (or column) of $\boldsymbol{A}$ is multiplied by a constant c, the new determinant is $c\det(\boldsymbol{A})$.
- If two rows or two columns of a matrix are interchanged, the sign of the determinant is changed.
- If two rows or two columns of a matrix are equal, then the determinant is zero. . . (see example on previous page).
- The determinant is unchanged by adding a multiple of some row to another row (the same is true for columns).
- A zero row or column implies a zero determinant.
- $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$

Inverse of a Matrix

- Let $\boldsymbol{A}$ be a square $(k \times k)$ matrix.
- The inverse of a matrix is the analog to scalar

$$a^{-1} = \frac{1}{a} \qquad \text{so} \qquad a^{-1}a = 1$$

- There may exist a matrix $\boldsymbol{B}_{(k \times k)}$ such that

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_k$$

.

- $\boldsymbol{B}$ is called the Inverse of $\boldsymbol{A}$ and is denoted as $\boldsymbol{B} = \boldsymbol{A}^{-1}$.
- The inverse exists *if and only if* the k columns (or rows) of $\boldsymbol{A}$ are linearly independent. i.e.,

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \ldots + c_k\boldsymbol{a}_k = 0$$

  where $c_1 = c_2 = \ldots = c_k = 0$.
- If $\boldsymbol{A}^{-1}$ exists, $\boldsymbol{A}$ is nonsingular.

  If $\boldsymbol{A}^{-1}$ does not exists, then $\boldsymbol{A}$ is singular.

Inverse of a Matrix

- A little example

$$\boldsymbol{A} = \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}, \qquad \boldsymbol{A}^{-1} = \frac{1}{10}\begin{pmatrix} -2 & 4 \\ 8 & -6 \end{pmatrix}$$

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \frac{1}{10}\begin{pmatrix} (-6+16) & (12-12) \\ (-8+8) & (16-6) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- Steps for computing the inverse of a matrix:
  - (1) Form a matrix of the same size as $\boldsymbol{A}$ containing the minors for all entries of $\boldsymbol{A}$ (i.e., for the $(i,j)^{th}$ entry, the minor is $\det(\boldsymbol{A}_{ij})$).
  - (2) Multiply the matrix of minors by $(-1)^{i+j}$ to produce the matrix of cofactors.
  - (3) Divide all entries in the cofactors matrix by $\det(\boldsymbol{A})$.
  - (4) The transpose of the matrix found in (3) gives $\boldsymbol{A}^{-1}$.

Example: Computing the Inverse

- The matrix $\boldsymbol{A}$ and it's determinant:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{and} \quad \det(\mathbf{A}) = -1$$

- (1) Matrix of minors:

$$\begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

(2) Matrix of cofactors (i.e., $(-1)^{i+j}$)

$$\begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{pmatrix}$$

- (3) Divide cofactors by $\det(\boldsymbol{A}) = -1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ -1 & -1 & 2 \\ -1 & 1 & -1 \end{pmatrix}$$

(4) Transpose gives $\boldsymbol{A}^{-1}$

$$\boldsymbol{A}^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix}$$

Inverse of a Matrix

- Check that answer is correct:

$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Computing inverse for 2×2 is easy:

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \boldsymbol{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

- Except for small matrices ($k = 2$ or 3), calculating the matrix inverse is (usually)difficult, so we'll let the computer do it for use.

- The exceptions are for diagonal and triangular matrices. . .

Inverse of a Diagonal and Triangular Matrices

Inverse of a Diagonal Matrix:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{pp} \end{pmatrix} \qquad \boldsymbol{A}^{-1} = \begin{pmatrix} 1/a_{11} & 0 & \ldots & 0 \\ 0 & 1/a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/a_{pp} \end{pmatrix}$$

Inverse of a Triangular matrix is a triangular matrix:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \ldots & a_{pp} \end{pmatrix} \qquad \boldsymbol{A}^{-1} = \begin{pmatrix} 1/a_{11} & 0 & \ldots & 0 \\ b_{21} & 1/a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \ldots & 1/a_{pp} \end{pmatrix}$$

Orthogonal Matrices

- $\boldsymbol{A}_{p\times p}$ is an orthogonal matrix if for all pairs of columns (rows)

$$\boldsymbol{a}_i'\boldsymbol{a}_k = \begin{cases} 0 & \text{for } i \neq k \\ 1 & \text{for } i = k \end{cases}$$

- If $\boldsymbol{A}$ is orthogonal, then $\boldsymbol{A}' = \boldsymbol{A}^{-1}$; that is, $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I} = \boldsymbol{A}\boldsymbol{A}'$ and $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} = \boldsymbol{A}'\boldsymbol{A}$.
- Because
  - $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I} \Longrightarrow$ Each column of $\boldsymbol{A}$ has length equal to 1 and all pairs of columns are orthogonal (i.e., perpendicular).
  - $\boldsymbol{A}\boldsymbol{A}' = \boldsymbol{I} \Longrightarrow$ Each row of $\boldsymbol{A}$ has length equal to 1 and all pairis of rows are orthogonal.
- If $\boldsymbol{A}$ is orthogonal, then $\det(\boldsymbol{A}) = \pm 1$.
- If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both orthogonal, the $\boldsymbol{A}\boldsymbol{B}$ is also orthogonal.
- Lengths of vectors don't change under orthogonal transformations; that is, let $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$; then

$$L^2_{\boldsymbol{y}} = \boldsymbol{y}'\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{x})'(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}'\boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}'\boldsymbol{x} = L^2_{\boldsymbol{x}}$$

Rank of a Matrix

- We can consider a matrix $\boldsymbol{A}_{n\times p}$ as a collection of column vectors and/or a collection of row vectors:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{r}'_1 \\ \boldsymbol{r}'_2 \\ \vdots \\ \boldsymbol{r}'_n \end{pmatrix} = (\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_p)$$

  where $\boldsymbol{r}'_i = (a_{i1}, a_{i2}, \ldots, a_{ip})$ and $\boldsymbol{c}'_j = (a_{1j}, a2j, \ldots, a_{nj})$.
- The Rank of a matrix is the maximum number of independent rows or columns in a matrix (row rank equals column rank).
- The rank of a matrix must be less than or equal to the smaller of n and p.
- Full Rank of a matrix means that the rank of the matrix equals the smaller of n or p.

- Outline

Motivation

Vectors

Matrices
- Matrices
- Matrix Operations & Definitions of Special Ones
- Equal Matrices
- Scalar Multiplication
- Matrix Addition
- Matrix Multiplication
- Matrix Multiplication
- Matrix Multiplication is Associative
- More Properties of Matrix Operations
- Square and Symmetric Matrices
- Trace of a Square Matrix
- Null and Diagonal Matrices
- Pre-Multiplication by a Diagonal Matrix
- Post-Multiplication by a Diagonal Matrix
- Identity Matrix
- Determinant of a Matrix
- Determinant of a $(p \times p)$ Matrix
- Example: Determinant of a (4×4) matrix
- More About Determinants
- Properties of Determinants
- Inverse of a Matrix

Rank Demonstration

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 5 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

Rows as vectors:

$$2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Linearly independent or dependent? What’s the row rank?

Columns as vectors

$$-2\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Linearly independent or dependent? What’s the column rank?

Usefulness Properties of Rank

- A and A' have the same rank.
- $A'A$, AA', and A have the same rank.
- The rank of a matrix is unchanged by a pre- or postmultiplication by a nonsingular matrix.
- The rank of a matrix is unchanged by what are called elementary row and column operations:
  - Interchange of two rows or two columns.
  - Multiplication or a row or a column by a scalar.
  - Addition of a row (or column) to another row (or column).

  This is true because any elementary operation can be represented by a premultiplication (if the operation is to be on rows) or a postmultiplication (if the operation is to be on columns) of a nonsingular matrix.

Quadratic Forms

- Suppose $\boldsymbol{A}_{n\times n}$ is symmetric and let $\boldsymbol{x}' = (x_1, \ldots, x_n)$. A quadratic form is defined by

$$\begin{aligned} \boldsymbol{x}'\mathbf{A}\boldsymbol{x} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_jx_k \\ &= a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{pp}x_n^2 \\ &\quad +2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + \cdots + 2a_{(n-1)n}x_{n-1}x_n \end{aligned}$$

- Why are quadratic forms important? One example,. . . What does this equal?

$$(x_1, x_2, \cdots, x_n)\begin{pmatrix} n/(1-n) & -1/n & \cdots & -1/n \\ -1/n & n/(1-n) & \cdots & -1/n \\ \vdots & \vdots & \ddots & \vdots \\ -1/n & -1/n & \cdots & n/(1-n) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$\sum_{j=1}^{n}(x_j - \bar{x})^2$

More about Quadratic Forms

- A symmetric matrix $\mathbf{A}$ (and associated quadratic form) is
  - positive definite (p.d.) if $\mathbf{x}'\mathbf{A}\boldsymbol{x} > 0$ for all $\boldsymbol{x} \neq \mathbf{0}$.
  - positive semi-definite (p.s.d) if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$ for all $\boldsymbol{x}$.
- Negative definite, negative semi-definite, and indefinite forms exist as well; however, correlation and covariance matrices (and $\boldsymbol{W}$)
  - Are at least positive semi-definite.
  - Satisfies the stronger condition of being positive definite if the vectors of the variables on which the correlation and covariance matrices is based, are linearly independent.
- Matrices that are positive definite have inverses (i.e., nonsingular).

First Pass at Eigenvalues and Eigenvectors

- A square matrix $\boldsymbol{A}_{p\times p}$ has an eigenvalue λ (a scalar) and a corresponding eigenvector $\boldsymbol{x}$ $(\neq \boldsymbol{0})$ if

$$\boldsymbol{Ax} = \lambda\boldsymbol{x} \qquad \text{or equivalently} \qquad (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$$

- Note:

$$\begin{aligned} \boldsymbol{Ax} &= \lambda\boldsymbol{x} \\ \boldsymbol{Ax} - \lambda\boldsymbol{x} &= \boldsymbol{0} \\ \boldsymbol{Ax} - \lambda\boldsymbol{Ix} &= \boldsymbol{0} \\ (\boldsymbol{Ax} - \lambda\boldsymbol{I})\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

- Usually we divide $\boldsymbol{x}$ by it's length $L_{\boldsymbol{x}}$ so that we get a normalized eigenvector; eigenvectors with length $= 1$,

$$\boldsymbol{e} = \frac{\boldsymbol{x}}{L_{\boldsymbol{x}}} \qquad \longrightarrow \qquad \boldsymbol{e}'\boldsymbol{e} = \left(\frac{\boldsymbol{x}'}{L_{\boldsymbol{x}}}\right)\left(\frac{\boldsymbol{x}}{L_{\boldsymbol{x}}}\right) = \frac{\boldsymbol{x}'\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = 1$$

Spectral Decomposition

- If $\boldsymbol{A}_{p\times p}$ is symmetric, then $\boldsymbol{A}$ has p pairs of eigenvalues and eigenvectors

$$\lambda_1, \boldsymbol{e}_1 \qquad \lambda_2, \boldsymbol{e}_2 \qquad \ldots \qquad \lambda_p, \boldsymbol{e}_p$$

- Matrix decomposition (very important)

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e}_1 \boldsymbol{e}_1' + \lambda_2 \boldsymbol{e}_2 \boldsymbol{e}_2' + \cdots + \lambda_p \boldsymbol{e}_p \boldsymbol{e}_p'$$

- We’ll cover this in more detail later, as well as some other topics (including SVD of rectangular matrices).

Random Vectors and Matrices

Mean Vectors, Covariance Matrices and Correlation Matrices

- Random vector is a vector whose elements are random variables.
- Random matrix is a matrix whose elements are random variables.
- Let $\boldsymbol{X} = \{X_{jk}\}_{n \times p}$ be a random matrix (i.e., data matrix).

  The expected value of $\boldsymbol{X}$ equals the expected values of it's entries (the random variables):

$$E(\boldsymbol{X}) = \begin{pmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1p}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{n1}) & E(X_{n2}) & \cdots & E(X_{np}) \end{pmatrix}$$

  assuming $E(X_{jk})$ exists for all (j, k).

Expected Values

- $E(X_{jk})$ equals the expected value or mean of random variable X_{jk}.

- If X_{jk} is discrete, then

$$E(X_{jk}) = \sum_{\text{all } x} x_{jk} p_{jk}(x_{jk})$$

where $p_{jk}(x_{jk})$ is the probability mass function of X_{jk}.

- If X_{jk} is continuous, then

$$E(X_{jk}) = \int_{-\infty}^{+\infty} x_{jk} f_{jk}(x_{jk}) dx_{jk}$$

where $f_{jk}(x_{jk})$ is the probability density function of X_{jk}.

In this class, our “dependent” variables will be continuous and we’ll use the multivariate normal distribution.

Expected Values (continued)

- Little example using a discrete variable
  - Let X_{11} be a Bernoulli random variable.
  - The probability mass function is

$$X_{11} = \begin{cases} 0 & \text{with probability } = 1 - \pi_{11} \\ 1 & \text{with probability } = \pi_{11} \end{cases}$$

  - The mean of X_{11} is

$$E(X_{11}) = 0(1 - \pi_{11}) + 1\pi_{11} = \pi_{11}$$

Here “all x” $= \{0, 1\}$.

Expected Values for Matrices

- Let $\boldsymbol{X}_{n\times p}$ and $\boldsymbol{Y}_{n\times p}$ be random matrices, then

$$E(\boldsymbol{X}+\boldsymbol{Y}) = E(\boldsymbol{X}) + E(\boldsymbol{Y})$$

- Also let $\boldsymbol{A}_{m\times n}$ and $\boldsymbol{B}_{p\times k}$ be matrices of constants, then

$$E(\boldsymbol{A}\boldsymbol{X}\boldsymbol{B}) = \boldsymbol{A}E(\boldsymbol{X})\boldsymbol{B}$$

- Preliminaries: Let $\boldsymbol{X}_{n\times 1}$ be a random vector where each element X_i is a random variable with

$$\mu_i = E(X_i) \qquad \text{and} \qquad \sigma_{ii} = E((X_i-\mu_i)^2)$$

for all $i = 1, \ldots, p$, and $\sqrt{\sigma_{ii}}$ equals the standard deviation.

- The covariance between X_i and X_k is

$$\sigma_{ik} = E((X_i-\mu_i)(X_k-\mu_k))$$

These expected values are computed using the joint probability mass function (if discrete) or the probability density function (if continuous) for X_i and X_k,

$$f_{ik}(x_i, x_k) \qquad \text{or} \qquad p_{ik}(x_i, x_k)$$

Implications of Statistical Independence

- Let's consider all p random variables simultaneously (i.e., the random vector)

$$\boldsymbol{X}' = (X_1, X_2, \cdots, X_p)$$

- Let $f(\boldsymbol{x})$ be the joint probability density (mass) function,

$$f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_p)$$

  In this class, $f(\boldsymbol{x})$ is the multivariate normal distribution.

- If X_i and X_k are statistically independent, then the joint probability function factors, e.g.,

$$f_{ik}(x_i, x_k) = f_i(x_i) f_k(x_k) \qquad \Longrightarrow \qquad \sigma_{ik} = 0$$

  Statistical independence implies $\sigma_{ik} = 0$.

- If all $\binom{p}{2}$ pairs of variables are statistically independent, then

$$f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_p) = f_1(x_1) f_2(x_2) \cdots f_p(x_p) \Longrightarrow \sigma_{ik} = 0 \quad \forall i, k$$

● Outline
Motivation
Vectors
Matrices
Random Vectors and Matrices
● Random Vectors and Matrices
● Expected Values
● Expected Values (continued)
● Expected Values for Matrices
● Implications of Statistical Independence
● Zero Covariance does not imply Independence
● Simplify Notation by Using Matrices
● Definition of Covariance Matrix
● Estimation of Population Quantities
● Estimation
● Correlation Matrices
● Correlation/Covariance Matrix Relationship
● Correlation/Covariance Matrix Relationship
● Estimation of Correlation Matrix
Useful Matrix Formulas
Linear Algebra
Matrix Software

Zero Covariance does not imply Independence

- In general, the reverse is not true; that is, a 0 covariance does not necessarily imply statistical independence.

  The exception is when $f(\boldsymbol{x})$ is multivariate normal.

- In sum,
  - Statistical independence implies 0 covariance.
  - 0 covariance does not imply statistical independence.
  - 0 covariance and multivariate normal implies statistical independence.

Simplify Notation by Using Matrices

$$E(\boldsymbol{X}) = \boldsymbol{\mu} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$$

and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

where $\sigma_{ik} = \sigma_{ki}$.

In sum, $\boldsymbol{\Sigma}$ is $p \times p$ symmetric matrix with variances along the diagonal and covariances on the off-diagonal.

Definition of Covariance Matrix

$$
\begin{aligned}
\boldsymbol{\Sigma} &= E((\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})') \\
&= E\left(\begin{pmatrix} (X_1-\mu_1) \\ (X_2-\mu_2) \\ \vdots \\ (X_p-\mu_p) \end{pmatrix}((X_1-\mu_1),(X_2-\mu_2),\cdots,(X_p-\mu_p))\right) \\
&= \begin{pmatrix}
E((X_1-\mu_1)^2) & E((X_1-\mu_1)(X_2-\mu_2)) & \cdots & E((X_1-\mu_1)(X_p-\mu_p)) \\
E((X_2-\mu_1)(X_1-\mu_1)) & E((X_2-\mu_2)^2) & \cdots & E((X_2-\mu_2)(X_p-\mu_p)) \\
\vdots & \vdots & \ddots & \vdots \\
E((X_p-\mu_p)(X_1-\mu_1)) & E((X_p-\mu_p)(X_2-\mu_2)) & \cdots & E((X_p-\mu_p)^2)
\end{pmatrix}
\end{aligned}
$$

Estimation of Population Quantities

The sample mean vector $\bar{x}$ is an estimator of $\boldsymbol{\mu}$:

$$\bar{\boldsymbol{x}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix}$$

The matrix of sample variances and covariances $\boldsymbol{S}$ is an estimator of $\boldsymbol{\Sigma}$:

$$\boldsymbol{S} = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix}$$

where $s_{ik} = \frac{1}{n}\sum_{j=1}^{n}(x_i - \bar{x}_i)(x_k - \bar{x}_k),$ so $s_{ik} = s_{ki}$

Estimation

- $\bar{\boldsymbol{x}}$ is the maximum likelihood estimator of $\boldsymbol{\mu}$.
- $\boldsymbol{S}$ is the maximum likelihood estimator of $\boldsymbol{\Sigma}$.
- Maximum likelihood estimators are denoted by $\hat{\ }$, i.e., $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$.
- $\hat{\boldsymbol{\Sigma}} = \boldsymbol{S}$ is a biased estimator of $\boldsymbol{\Sigma}$. For an unbiased one, divide sum of squared deviations by $(n-1)$.
- The sum of squared deviations and sum of cross-product deviations

  $$\boldsymbol{W} = \{w_{ik}\}_{p\times p}$$

  where $w_{ik} = \sum_{j=1}^{n}(x_{ji} - \bar{x}_i)(x_{jk} - \bar{x}_k)$, which are the numerators of the variances and covariances.

Correlation Matrices

The population (Pearson product-moment) correlation coefficient is

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$$

which measures the amount of linear association between the random variables X_i and X_k.

The population correlation matrix:

$$\boldsymbol{P} =$$

$$\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \sigma_{11}/\sqrt{\sigma_{11}\sigma_{11}} & \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}} & \cdots & \sigma_{1p}/\sqrt{\sigma_{11}\sigma_{pp}} \\ \sigma_{21}/\sqrt{\sigma_{22}\sigma_{11}} & \sigma_{22}/\sqrt{\sigma_{22}\sigma_{22}} & \cdots & \sigma_{2p}/\sqrt{\sigma_{22}\sigma_{pp}} \\ \vdots & \vdots & \ldots & \vdots \\ \sigma_{p1}/\sqrt{\sigma_{pp}\sigma_{11}} & \sigma_{p2}/\sqrt{\sigma_{pp}\sigma_{22}} & \cdots & \sigma_{pp}/\sqrt{\sigma_{pp}\sigma_{pp}} \end{pmatrix}$$

Correlation/Covariance Matrix Relationship

- Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices
- Random Vectors and Matrices
- Expected Values
- Expected Values (continued)
- Expected Values for Matrices
- Implications of Statistical Independence
- Zero Covariance does not imply Independence
- Simplify Notation by Using Matrices
- Definition of Covariance Matrix
- Estimation of Population Quantities
- Estimation
- Correlation Matrices
- Correlation/Covariance Matrix Relationship
- Correlation/Covariance Matrix Relationship
- Estimation of Correlation Matrix

Useful Matrix Formulas

Linear Algebra

Matrix Software

in matrix terms. . .

Define

$$\boldsymbol{D}^{-1/2}_{p\times p} = \begin{pmatrix} 1/\sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{\sigma_{pp}} \end{pmatrix}$$

Recall:

- Pre-multiplication of a matrix by a diagonal matrix $\longrightarrow$ multiply the j^{th} row of the matrix by the d_{jj} (j^{th} diagonal elements).
- Post-multiplication of a matrix by a diagonal matrix $\longrightarrow$ multiply the k^{th} column of the matrix by the k^{th} diagonal element, d_{kk}.

Correlation/Covariance Matrix Relationship

So

$$\boldsymbol{P} = \boldsymbol{D}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{D}^{-1/2} \qquad \text{and} \qquad \boldsymbol{\Sigma} = \boldsymbol{D}^{1/2}\boldsymbol{P}\boldsymbol{D}^{1/2}$$

where $\boldsymbol{D}^{1/2} = \text{diag}(\sqrt{\sigma_{ii}})$.
e.g.,

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 3 & 9 \end{pmatrix} \qquad \boldsymbol{D}^{-1/2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

$$\begin{aligned} \boldsymbol{P} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 3 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1/2 & 2/6 \\ 1/2 & 1 & 1/2 \\ 2/3 & 1/5 & 1 \end{pmatrix} \end{aligned}$$

Estimation of Correlation Matrix

The sample correlation matrix $\boldsymbol{R}$ estimates $\boldsymbol{P}$,

$$\boldsymbol{R} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix}$$

where $r_{ik} = s_{ik}/(\sqrt{s_{ii}}\sqrt{s_{kk}})$.

The relationship between $\boldsymbol{S}$ and $\boldsymbol{R}$ is the same as the one between $\boldsymbol{\Sigma}$ and $\boldsymbol{P}$.

Let $\boldsymbol{D}_s^{-1/2} = \text{diag}(\sqrt{s_{ii}})$, then

$$\boldsymbol{S} = \boldsymbol{D}_s^{1/2}\boldsymbol{R}\boldsymbol{D}_s^{1/2} \qquad \text{and} \qquad \boldsymbol{R} = \boldsymbol{D}_s^{-1/2}\boldsymbol{S}\boldsymbol{D}_s^{-1/2}$$

Useful Matrix Formulas: Mean vector

$(n \times p)$ Data Matrix $\boldsymbol{X}$

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix}$$

$(p \times 1)$ Mean vector $\bar{\boldsymbol{x}}$

$$\bar{\boldsymbol{x}} = \frac{1}{n}\boldsymbol{X}'\mathbf{1}_n = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix} = \frac{1}{n}\begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \dots & x_{np} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

or as a row vector

$$\bar{\boldsymbol{x}}' = \frac{1}{n}\mathbf{1}_n'\boldsymbol{X}$$

Matrix of Deviations from Means

Note:

$$\mathbf{1}_n\bar{\boldsymbol{x}}' = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p) = \begin{pmatrix} \bar{x}_1 & \bar{x}_2 & \ldots & \bar{x}_p \\ \bar{x}_1 & \bar{x}_2 & \ldots & \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_1 & \bar{x}_2 & \ldots & \bar{x}_p \end{pmatrix}$$

So,

$$\boldsymbol{X} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X} = \begin{pmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \ldots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & \ldots & x_{2p} - \bar{x}_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & \ldots & x_{np} - \bar{x}_p \end{pmatrix}$$

Covariance Matrix

● Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices

Useful Matrix Formulas

● Useful Matrix Formulas: Mean vector

● Matrix of Deviations from Means

● Covariance Matrix

● Correlation Matrix

Matrix Software

$(p \times p)$ Variance/Covariance Matrix $\boldsymbol{S}$

$$\begin{aligned}
\boldsymbol{S} &= \left(\frac{1}{n-1}\right)\left(\boldsymbol{X} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X}\right)'\left(\boldsymbol{X} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\boldsymbol{X}\right) \\
&= \left(\frac{1}{n-1}\right)\boldsymbol{X}'\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)'\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\boldsymbol{X} \\
&= \left(\frac{1}{n-1}\right)\boldsymbol{X}'\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\boldsymbol{X}
\end{aligned}$$

Since

$$\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)'\left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right) = \left(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)$$

Also,

$$\boldsymbol{W} = (n-1)\boldsymbol{S} = \boldsymbol{X}'(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\boldsymbol{X}$$

Correlation Matrix

- Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices

Useful Matrix Formulas
- Useful Matrix Formulas: Mean vector
- Matrix of Deviations from Means
- Covariance Matrix
- Correlation Matrix

Matrix Software

Define:

$$\boldsymbol{D}_{p\times p} = \begin{pmatrix} s_{11} & 0 & \dots & 0 \\ 0 & s_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & s_{pp} \end{pmatrix}$$

and

$$\boldsymbol{D}^{-1/2}_{p\times p} = \begin{pmatrix} 1/\sqrt{s_{11}} & 0 & \dots & 0 \\ 0 & 1/\sqrt{s_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\sqrt{s_{pp}} \end{pmatrix}$$

$(p \times p)$ Correlation Matrix $\boldsymbol{R}$

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{D}^{-1/2}\boldsymbol{S}\boldsymbol{D}^{-1/2} \\ &= \left(\frac{1}{n-1}\right)\boldsymbol{D}^{-1/2}\boldsymbol{X}'\left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{1}_n\boldsymbol{1}_n'\right)\boldsymbol{X}\boldsymbol{D}^{-1/2} \end{aligned}$$

And

$$\boldsymbol{S} = \boldsymbol{D}^{1/2}\boldsymbol{R}\boldsymbol{D}^{1/2}$$

Matrix Software: SAS/IML

- Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices

Useful Matrix Formulas

Matrix Software
- Matrix Software: SAS/IML
- SAS/IML (continued)
- MATLAB
- R (SPlus)
- R (continued)

```
proc iml;
* Create matrices;
  A =   {1   2   3,
          5   1   2};

  B =   {2   0,
          4   3,
         -2   1};

* Matrix addition;
E = A + B‘;
print E;

* Matrix Multiplication;
AB=A*B;
BA=B*A;
print ’Multiplication is (usually) not commutative’,
           ’AB= ’ AB,
           ’BA= ’ BA;
```

SAS/IML (continued)

```
* Determinant of a matrix;
 A =    {4   3   2    1,
          2   5   1    0,
          1   3   2    1,
          3   2   1   4};
detA = det(A);
print detA;

* Inverse of a matrix;
invA = inv(A);
check1 = invA*A;
check2 = A*invA;

* Trace of a matrix;
trA = trace(A);

print invA, check1, check2, trA;
```

MATLAB

```
A =   [1   2   3;
       5   1   2]
B =   [2   0;
       4   3;
      -2   1]
E = A + B’
AB=A*B
BA=B*A
A =   [4   3   2   1;
       2   5   1   0;
       1   3   2   1;
       3   2   1   4]
detA = det(A)
invA = inv(A)
invA*A;
A*invA;
trace(A)
```

R (SPlus)

- Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices

Useful Matrix Formulas

Matrix Software
- Matrix Software: SAS/IML
- SAS/IML (continued)
- MATLAB
- R (SPlus)
- R (continued)

```
# Example of R code for simple matrix operations
#
# Create matrices

A <- matrix(c(1,5, 2,1, 3,2), nrow=2, ncol=3)
B <- matrix(c(2, 4, -2, 0, 3, 1), nrow=3, ncol=2)

# Matrix addition (A plus transpose of B)
# Matrix multiplication

AB <- A %*% B
BA <- B %*% A

# Note that the following does element by
# element multiplication

AtB <- A * t(B)
```

R (continued)

- Outline

Motivation

Vectors

Matrices

Random Vectors and Matrices

Useful Matrix Formulas

Matrix Software
- Matrix Software: SAS/IML
- SAS/IML (continued)
- MATLAB
- R (SPlus)
- R (continued)

```
# Determinant of a matrix

A <- matrix(c(4, 2, 1, 3,
              3, 5, 3, 2,
              2, 1, 2, 1,
              1, 0, 1, 4),
              nrow=4, ncol=4)

det <- det(A)

# inverse of a matrix
invA <- solve(A)

# Check
AinvA <- A %*% invA
invAA <- invA %*% A
```