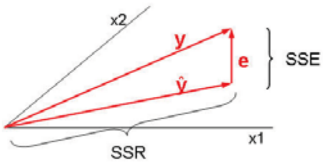
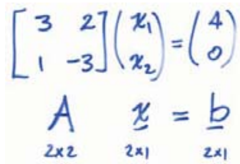
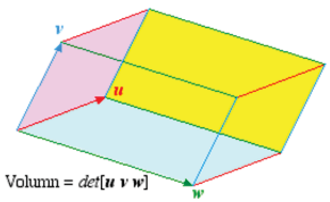




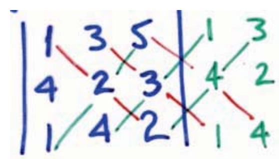
$$\mathbf{R}\begin{pmatrix}\mathbf{x}\\ \hline \mathbf{y}\end{pmatrix}=\left[\begin{array}{c|c}\mathbf{R}_{XX} & \mathbf{R}_{XY}\\ \hline \mathbf{R}_{YX} & \mathbf{R}_{YY}\end{array}\right]$$

Vectors & Matrices

with statistical applications



Michael Friendly
Psychology 6140



Why learn matrix algebra?

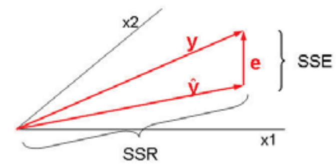
- Simple way to express linear combinations of variables and general solutions of equations.

$$\begin{cases}\mathbf{a}'\mathbf{x}=a_1x_1+a_2x_2+a_3x_3\\ \mathbf{A}_{n\times n}\mathbf{x}_{n\times 1}=\mathbf{b}_{n\times 1}\Rightarrow \mathbf{x}=\mathbf{A}^{-1}\mathbf{b}\end{cases}$$

- Linear statistical models (regression, anova) generalize to any # of predictors & responses.

$$\begin{cases}\hat{y}_i=\beta_0+\beta_1x_1+\beta_2x_2\\ \hat{\mathbf{y}}=\mathbf{X}\boldsymbol{\beta}\\ \hat{\mathbf{Y}}=\mathbf{X}\mathbf{B}\end{cases}$$

- Strong relations between algebra, geometry & statistical concepts

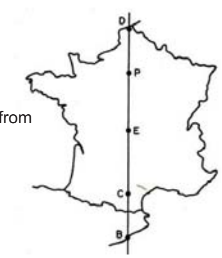


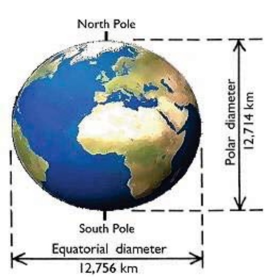
Goal: a *reading* knowledge of matrix expressions to aid in understanding statistical concepts.

Brief history of linear algebra

- Ideas first arose in relation to solving systems of equations in astronomy & geodesy (1700s)
  - Determining the “shape of the earth” from measures of latitude and longitude (3 eqn., 3 unknowns)
  - Calculating the orbits of planets, e.g., Saturn, Jupiter (6 eqn., 6 unknowns)

Arc lengths measured from Dunkirk to Barcelona



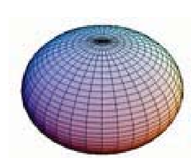


Pierre-Louis Moreau de Maupertius

“*The man who flattened the earth*”

(Portrait from 1739)

His crowning glory was a journey to Lapland, making measures of the length of 1º of latitude, and showing that they were smaller near the poles than at the equator.



Brief history of linear algebra

- By ~ 1800, Gauss developed "Gaussian elimination" to solve such problems, and "least squares" to deal with fallible measurements
- Still required proper notation & algebra ($\mathbf{A}_{m \times n}$)
  - 1848: J.J. Sylvester introduced "matrix" (Latin: womb) for array of numbers, with a *single symbol*.
  - 1855: Arthur Cayley defined matrix multiplication in relation to systems of equations
  - 1858: Cayley develops algebra, including inverse, $\mathbf{A}^{-1}$
- Now, there was a **general** notation for solving *m* equations in *n* unknowns!

Vectors & matrices

- A *matrix* is a rectangular array of numbers, with *r* rows and *c* columns.

$$\mathbf{A}_{3\times 2} = \begin{bmatrix} 12 & 3 \\ 15 & 0 \\ 7 & -1 \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = (a_{ij}),\ {}_{j=1,2,\ldots c}^{i=1,2,\ldots r}$$

$$\mathbf{B}_{2\times 3} = \begin{pmatrix} 1 & 7 & -3 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

Transpose operation: $\mathbf{A}' \equiv \mathbf{A}^{\mathsf{T}} = [a_{ji}]$, interchanges rows and columns

$$\mathbf{B}'_{2\times 3} = \begin{pmatrix} 1 & 2 \\ 7 & 4 \\ -3 & 6 \end{pmatrix}$$

Vectors & matrices

- A *vector* is just a one column matrix
- Sometimes written in transposed (row) form to save space.

$$\mathbf{y}_{3\times 1} = \begin{pmatrix} 6 \\ 7 \\ 12 \end{pmatrix}$$

$$\mathbf{y}'_{1\times 3} \equiv \mathbf{y}^{T}{}_{1\times 3} = \begin{pmatrix} 6 & 7 & 12 \end{pmatrix}$$

$$\mathbf{y}_{3\times 1} = \begin{pmatrix} 6 & 7 & 12 \end{pmatrix}'$$

All of these forms define **y** as a 3 x 1 column vector

Special vectors & matrices

unit vector: $\underline{1}_n \equiv \underline{j}_n = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_n$

zero vector: $\underline{0}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_n$

contrast vectors: $\sum^{n} c_i = 0$

$$\mathbf{c}'_1 = \begin{pmatrix} 1 & 1 & -1 & -1 \end{pmatrix}$$

$$\mathbf{c}'_2 = \begin{pmatrix} 3 & -1 & -1 & -1 \end{pmatrix}$$

Square matrix: $\mathbf{A}_{n \times n}$: same # rows/cols

$$\mathbf{A}_{2\times 2} = \begin{bmatrix} 2 & 10 \\ 11 & 9 \end{bmatrix} \qquad \mathbf{B}_{3\times 3} = \begin{bmatrix} 9 & 7 & 1 \\ 3 & 3 & 5 \\ 1 & 9 & 4 \end{bmatrix}$$

Symmetric matrix: $\mathbf{A} = \mathbf{A}^{\mathsf{T}}$, or $a_{ij}=a_{ji}$

$$\mathbf{A}_{2\times 2} = \begin{bmatrix} 2 & 10 \\ 10 & 9 \end{bmatrix} \qquad \mathbf{B}_{3\times 3} = \begin{bmatrix} 9 & 7 & 1 \\ 7 & 3 & 5 \\ 1 & 5 & 4 \end{bmatrix}$$

Diagonal matrix: $a_{ij} = 0$ for $i \neq j$

$$\mathbf{D}_{2\times 2} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{D}_{3\times 3} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Special vectors & matrices

Identity matrix: diagonal w/ $a_{ii} = 1$

$$\mathbf{I}_{2\times2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \mathbf{I}_{3\times3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Why: acts like 1 in multiplication–

A I = A

Unit matrix: all $a_{ij} = 1$

$$\mathbf{J}_{3\times2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{j}_3 & \mathbf{j}_3 \end{bmatrix}$$

Why: convenient way to sum vectors & matrices

$\mathbf{a}^\mathsf{T}\mathbf{j} = \sum a_i$

Zero matrix: all $a_{ij} = 0$

$$\mathbf{0}_{2\times3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Why: acts like 0 in addition—

A + 0 = A

A - B = 0 → A = B

Operations on vectors & matrices

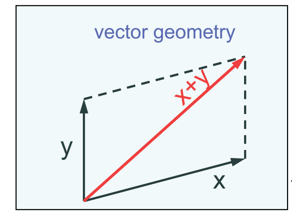
Addition & subtraction: add corresponding elements. Must have same shape

$$\mathbf{a}_{3\times1} + \mathbf{b}_{3\times1} = (a_i + b_i) = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \qquad \mathbf{A}_{3\times2} + \mathbf{B}_{3\times2} = \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 10 & 2 \\ 4 & 6 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 5 & 3 \\ 4 & 4 \end{bmatrix} \Rightarrow \mathbf{A} + \mathbf{B} = \begin{bmatrix} 15 & 5 \\ 8 & 10 \end{bmatrix} \quad \mathbf{A} - \mathbf{B} = \begin{bmatrix} 5 & -1 \\ 0 & 2 \end{bmatrix}$$

Properties: same as for scalars– order doesn't matter

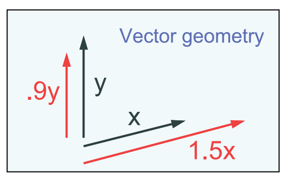
- Commutative: **A + B = B + A**
- Associative: **A + (B + C) = (A + B) + C**



Operations on vectors & matrices

Scalar multiplication: multiply each element by a scalar.

$$k\mathbf{a}_{n\times1} = (ka_i) = \begin{pmatrix} ka_1 \\ \vdots \\ ka_n \end{pmatrix} \qquad c\mathbf{A}_{m\times n} = \begin{bmatrix} ca_{ij} \end{bmatrix} = \begin{pmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{pmatrix}$$



$$3\begin{pmatrix} 1 \\ 5 \\ 10 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \\ 30 \end{pmatrix} \qquad 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} \qquad \lambda\mathbf{I}_{n\times n} = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

$$R - \lambda\mathbf{I} = \begin{bmatrix} 1-\lambda & r_{12} & r_{13} \\ r_{21} & 1-\lambda & r_{23} \\ r_{31} & r_{32} & 1-\lambda \end{bmatrix}$$

Partitioned matrices

Def^n : A partitioned matrix has its rows & columns divided into sub-matrices

$$\mathbf{A}_{4\times3} = \left[\begin{array}{cc:c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hdashline 7 & 8 & 9 \\ 10 & 11 & 12 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

$$\mathbf{A}_{21} = \begin{bmatrix} 7 & 8 \\ 10 & 11 \end{bmatrix}$$

Statistical examples:

$$\mathbf{R}\left(\begin{array}{c} \mathbf{x} \\ \hdashline \mathbf{y} \end{array}\right) = \left[\begin{array}{c:c} \mathbf{R}_{XX} & \mathbf{R}_{XY} \\ \hdashline \mathbf{R}_{YX} & \mathbf{R}_{YY} \end{array}\right] \qquad (\mathbf{X} \,\vdots\, \mathbf{y})'(\mathbf{X} \,\vdots\, \mathbf{y}) = \left[\begin{array}{c:c} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{y} \\ \hdashline \mathbf{y}'\mathbf{X} & \mathbf{y}'\mathbf{y} \end{array}\right]$$

Partitioned matrices

Addition and subtraction is defined for partitioned matrices if all submatrices in corresponding positions are of the same size and shape

$$\left[\begin{array}{cc|c}1 & 2 & 3\\ 4 & 5 & 6\\ \hline 7 & 8 & 9\end{array}\right] + \left[\begin{array}{cc|c}2 & 2 & 0\\ 1 & 0 & -1\\ \hline 0 & 2 & -1\end{array}\right] = \left[\begin{array}{cc|c}3 & 4 & 4\\ 5 & 5 & 5\\ \hline 7 & 10 & 8\end{array}\right]$$

Symbolically,

$$\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} + \begin{bmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{11}+\mathbf{B}_{11} & \mathbf{A}_{12}+\mathbf{B}_{12}\\ \mathbf{A}_{21}+\mathbf{B}_{21} & \mathbf{A}_{22}+\mathbf{B}_{22}\end{bmatrix}$$

Vector & matrix multiplication

Vector x vector
(inner product)

$$\underset{1\times n}{\underline{a}'}\ \underset{n\times 1}{\underline{x}} = (a_1\ a_2 \cdots a_n)\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$$

$$= a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

$$= \sum_{i=1}^{n} a_i x_i$$ sum of products of corresponding elements

eg: $(1\ \ -2\ \ 1)\begin{pmatrix}12\\ 20\\ 40\end{pmatrix} = 1\cdot 12 + (-2)\cdot 20 + 1\cdot 40 = 12$

NB: $\underline{a}'\underline{x} \equiv \underline{x}'\underline{a}$

Note that inner dimensions must match!

Vector & matrix multiplication

special cases

(a) $\underline{y}'\underline{y} = (y_1,\dots,y_n)\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix} = y_1^2 + y_2^2 + \cdots y_n^2 = \sum_{i=1}^{n} y_i^2$

(b) $\underline{1}'\underline{y} = (1\ 1 \cdots 1)\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix} = y_1 + y_2 + \cdots y_n = \sum y_i$

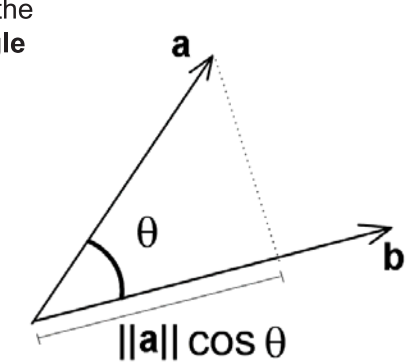
$= \underline{y}'\underline{1}$

(c) $\left.\begin{array}{l}\underline{a}'\underline{0} = 0\\ \underline{0}'\underline{a} = 0\end{array}\right\}$ $(1\ \ 3\ \ 5)\begin{pmatrix}0\\ 0\\ 0\end{pmatrix} = 0$

Geometry of vector products

- In a geometric representation, the scalar product relates to the **angle** between 2 vectors:

$$\mathbf{a}'\mathbf{b} = \|\mathbf{a}\|\cdot\|\mathbf{b}\|\cdot\cos\theta$$



- **Orthogonal** vectors (θ=90) have the property that **a′ b** = 0

$$\mathbf{a}'\mathbf{b} = (1\ \ 1\ \ 1)\begin{pmatrix}1\\ -2\\ 1\end{pmatrix} = 0$$

- **Correlation** (= cos θ) = $\dfrac{\mathbf{x}'\mathbf{y}}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}$

Matrix product

The matrix product, **A B**, is defined only if

the # of columns of A = # of rows of B

Then, **A** and **B** are conformable for multiplication

$$\underbrace{\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}}_{(2\times 3)}\underbrace{\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}}_{(3\times 3)} \quad \checkmark$$

$$\underbrace{\begin{bmatrix}1&1&1\\2&2&2\\3&3&3\end{bmatrix}}_{(3\times 3)}\underbrace{\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}}_{(2\times 3)}$$

Not conformable

Matrix product

Algebraic view:

matrix x matrix

Let $\underset{r\times c}{A} = [a_{ik}]$ $\quad i=1,\cdots r \quad k=1,\cdots c$

$\underset{c\times s}{B} = [b_{kj}]$ $\quad k=1,\cdots c \quad j=1,\cdots s$

Then

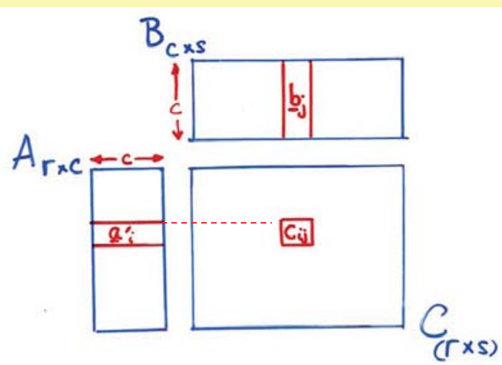
$$\underset{r\times c}{A}\cdot\underset{c\times s}{B} = \underset{r\times s}{C} = [c_{ij}] = [\underline{a}_i'\,\underline{b}_j]$$

product

rows of A · cols of B

Matrix product

Diagram view:



or:

$$c_{ij} = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{ic}\end{pmatrix}\begin{pmatrix} b_{1j}\\ b_{2j}\\ \vdots \\ b_{cj}\end{pmatrix} = \sum_{k=1}^{c} a_{ik}b_{kj}$$

Matrix product: examples

e.g.

$$\underset{A}{\begin{pmatrix}1&1&2\\4&0&-1\end{pmatrix}}\underset{B}{\begin{pmatrix}2&0\\1&1\\0&3\end{pmatrix}} = \begin{pmatrix}2+1+0 & 0+1+6\\ 8+0+0 & 0+0-3\end{pmatrix} = \begin{pmatrix}3&7\\8&-3\end{pmatrix}$$

2x3 ✓ 3x2 $\qquad$ 2x2

$$\underset{A}{\begin{pmatrix}1&2\\3&4\end{pmatrix}}\underset{B}{\begin{pmatrix}0&3\\2&1\end{pmatrix}} = \begin{pmatrix}4&5\\8&13\end{pmatrix}$$

NB $AB \neq BA$

$$\underset{B}{\begin{pmatrix}0&3\\2&1\end{pmatrix}}\underset{A}{\begin{pmatrix}1&2\\3&4\end{pmatrix}} = \begin{pmatrix}9&12\\5&8\end{pmatrix}$$

Why multiply like this?

To express systems of linear equations:

$$\begin{aligned} 3x_1 + 2x_2 &= 4 \\ x_1 - 3x_2 &= 0 \end{aligned} \iff \begin{bmatrix} 3 & 2 \\ 1 & -3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$$

$$\underset{2\times 2}{A} \quad \underset{2\times 1}{\underline{x}} = \underset{2\times 1}{\underline{b}}$$

In general:

$$\left.\begin{matrix} m \text{ equations} \\ n \text{ unknowns} \end{matrix}\right\} \rightarrow \underset{(m\times n)}{A} \quad \underset{(n\times 1)}{\underline{x}} = \underset{(m\times 1)}{b}$$

Solution: $\mathbf{x} = \mathbf{A}^{-1}\,\mathbf{b}$ when $\mathbf{A}^{-1}$ exists (m=n, eqn. independent)

Properties of matrix multiplication

1. Associative $\quad \mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
2. Distributive $\quad (\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$
   $\quad \mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
3. NOT commutative (in general) $\quad \mathbf{AB} \neq \mathbf{BA}$
4. Identity $\quad \underset{r\times c}{\mathbf{A}}\;\underset{c\times c}{\mathbf{I}} = \underset{r\times r}{\mathbf{I}}\;\underset{r\times c}{\mathbf{A}} = \mathbf{A}$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Properties of matrix multiplication

5. Zero $\quad \mathbf{A}_{r\times c}\;\mathbf{0}_{c\times s} = \mathbf{0}_{r\times s}$
6. Transpose of a product $\quad (\mathbf{AB})^T = \mathbf{B}^T\,\mathbf{A}^T$
   $\quad (\mathbf{AB}\cdots\mathbf{Z})^T = \mathbf{Z}^T \cdots \mathbf{B}^T\,\mathbf{A}^T$

All of these properties are analogous to ordinary (scalar) algebra, except for (3) and (6). Why?

Matrix powers

For a **square** matrix, $A_{(n \times n)}$:

$$\mathbf{A}^2 = \mathbf{A}\;\mathbf{A}$$

$$\mathbf{A}^3 = \mathbf{A}\;\mathbf{A}\;\mathbf{A} = \mathbf{A}^2\mathbf{A}$$ etc, for A^p

e.g., $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$$

In applications (e.g., MAP II-1), matrix powers provide a simple way to compute paths through a network, represented by (0/1) values in a matrix.

Matrix powers

Square roots too:

If $\mathbf{B}^2 = \mathbf{A}$, then $\mathbf{B}$ is also the square root of $\mathbf{A}$, i.e., $\mathbf{B} = \mathbf{A}^{1/2}$

e.g., $$\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}^2 = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 0 \\ 0 & 9 \end{pmatrix} = \mathbf{A}$$

so, $$\begin{pmatrix} 16 & 0 \\ 0 & 9 \end{pmatrix}^{1/2} = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} = \mathbf{B} = \mathbf{A}^{1/2}$$

The idea of the "square root of a matrix" was fundamental in the development of factor analysis, where Thurstone defined factors as

$$\boxed{\mathbf{R}} \approx \boxed{\Lambda}\ \boxed{\Lambda'}$$

Vectors & matrices in regression

The general linear regression model,

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i$$

has the following form in terms of vectors and matrices:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

or,

$$\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times(p+1)}\boldsymbol{\beta}_{(p+1)\times 1} + \boldsymbol{\varepsilon}_{n\times 1}$$

Matrix products in regression

All calculations are based on the sums and sums of squares from the following matrix products (shown for p=1 predictor):

$$y'y = (y_1, y_2, \dots y_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \underbrace{\sum_{i=1}^{n} y_i^2}_{1\times 1}$$

$$X'X = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & & x_n \end{bmatrix}\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \underbrace{\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}}_{2\times 2}$$

$$X'y = \underbrace{\begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & & x_n \end{bmatrix}}_{2\times n}\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{n\times 1} = \underbrace{\begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}}_{2\times 1}$$

We can represent these all with partitioned matrices:

$$(\mathbf{X} \mid \mathbf{y})'(\mathbf{X} \mid \mathbf{y}) = \left[\begin{array}{c|c} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{y} \\ \hline \mathbf{y}'\mathbf{X} & \mathbf{y}'\mathbf{y} \end{array}\right]$$

Linear combinations of vectors

Given vectors $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$, etc (same length)

a linear combination is a weighted sum

of the form

$$a(\underline{x}_1) + b(\underline{x}_2) + c(\underline{x}_3)$$

a, b, c scalars

e.g. $3\underline{x}_1 + 2\underline{x}_2 - 7\underline{x}_3$

Linear combinations of vectors

If $\underline{x}_1 = \begin{pmatrix}3\\1\end{pmatrix}$, $\underline{x}_2 = \begin{pmatrix}1\\2\end{pmatrix}$, $\underline{x}_3 = \begin{pmatrix}0\\1\end{pmatrix}$

then $3\underline{x}_1 + 2\underline{x}_2 - 7\underline{x}_3$

$= 3\begin{pmatrix}3\\1\end{pmatrix} + 2\begin{pmatrix}1\\2\end{pmatrix} - 7\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}11\\0\end{pmatrix}$, another vector

Linear independence

- A set of vectors, $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_n$ is linearly dependent if:
  1. One $\mathbf{x}_i$ can be expressed as a linear combination of the others; or, equivalently:
  2. There are some scalars, $a_1, a_2, \ldots a_n$, not all zero, such that

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \ldots + a_n\mathbf{x}_n = \mathbf{0} = \begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$$

Otherwise, the vectors are linearly independent.

Why: linear independence → idea of rank of a matrix

Linear independence: example

eg

$$\begin{array}{cccc} & \text{Verbal} & \text{Math} & \text{Composite} \\ & x_1 & x_2 & x_3 = 2\underline{x}_1 + \underline{x}_2 \end{array}$$

$$X = \begin{bmatrix} 10 & 12 & 32 \\ 8 & 4 & 20 \\ 5 & 10 & 20 \\ 15 & 5 & 35 \end{bmatrix}$$

$\underline{x}_1, \underline{x}_2, \underline{x}_3$ are linearly dependent

(a) since $\underline{x}_3 = 2\underline{x}_1 + \underline{x}_2$

(b) since $2\underline{x}_1 + \underline{x}_2 - \underline{x}_3 = \underline{0}$

$\therefore$ $\underline{x}_3$ provides no new information not provided by $\underline{x}_1$ & $\underline{x}_2$ already ($\underline{x}_3$ is redundant, given $\underline{x}_1$ & $\underline{x}_2$)

Rank of a matrix

The idea of rank of a matrix (or set of vectors) is a fundamental idea in matrix algebra and statistical applications.

- **Def**: rank($\mathbf{A}$) Ξ r ($\mathbf{A}$) = # of linearly independent rows (or columns) of $\mathbf{A}_{r \times c}$
- Properties:
  - # linearly independent rows = # linearly independent columns
  - r ($\mathbf{A}$) ≤ min(r, c) – rank never greater than smaller dimension
  - r ($\mathbf{A}$ $\mathbf{B}$) = min[r($\mathbf{A}$), r($\mathbf{B}$)] – rank of product = smaller of separate ranks

- **Geometric idea**: rank = # of dimensions (of a vector space)
- **Statistical idea**: rank = degrees of freedom

  = # of linearly independent variables

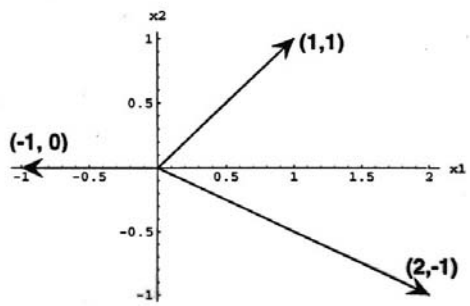
Rank and Dimensionality of Vector Spaces

Row space of a matrix: vector space spanned by rows

Column space of a matrix: vector space spanned by columns

$$A_{(3\times2)} = \begin{array}{cc} & \begin{array}{cc} x_1 & x_2 \end{array} \\ & \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & -1 \end{pmatrix} \begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \end{array}$$

Row space



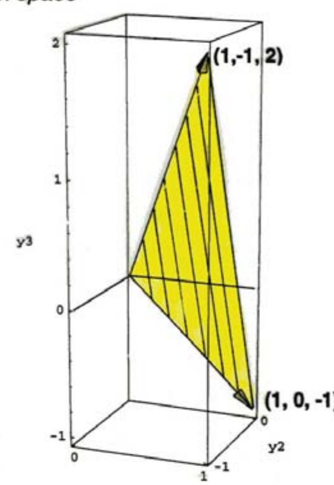
The rows of A define 3 vectors in 2-space, a plane. The columns define 2 vectors in 3-space-- also a plane. Hence, rank(A) = 2.

Rank and Dimensionality of Vector Spaces

rank(A)	*Row/Col space*
0	point
1	line
2	plane
3	volume
4+	hypervolume

$$A_{(3\times2)} = \begin{array}{cc} & \begin{array}{cc} x_1 & x_2 \end{array} \\ & \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 2 & -1 \end{pmatrix} \begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \end{array}$$

Column space



Matrix inverse: A^{-1}

- Inverse of a number

In scalar algebra, division (inverse of multiplication) is essential for solving equations.

e.g. $4x = 8 \rightarrow x = 8/4 = 2$

Equally, we can regard this as multiplying both sides by the inverse of the constant:

$$4x = 8 \rightarrow \tfrac{1}{4}\cdot 4x = \tfrac{1}{4}\cdot 8$$
$$1x = 2$$

Matrix inverse: A^{-1}

- Inverse of a matrix

Although division is not defined for matrices, most square matrices have a matrix inverse, which plays a similar role in solving equations

The inverse of an $n\times n$ matrix A is an $n\times n$ matrix A^{-1} such that

$$A A^{-1} = A^{-1}A = I_{(n\times n)}$$

e.g. $D = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$ - then $D^{-1} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/3 \end{pmatrix}$

because $\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1/4 & 0 \\ 0 & 1/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$

or $D \quad D^{-1} = I$

Matrix inverse: basic properties

- If an inverse, $\mathbf{A}^{-1}$ exists, it is unique
- No inverse exists if $\mathbf{A}_{n \times n} = \mathbf{0}$ (i.e., r(A)=0) or, in general, if r(A) < n
  - $\rightarrow$ **A** is 'singular'
  - $\rightarrow$ det(**A**) = |**A**| = 0
- Ordinary inverse defined only for square, non-singular matrices
  - Can also define a 'generalized inverse,' $\mathbf{A}^{-}$, such that $\mathbf{A}\,\mathbf{A}^{-}\,\mathbf{A} = \mathbf{A}$ and $\mathbf{A}^{-}\,\mathbf{A}\,\mathbf{A}^{-} = \mathbf{A}^{-}$

Matrix inverse: 2 x 2

The inverse of a 2 x 2 matrix is easy to calculate:

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

e.g.,

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ -1 & 4 \end{bmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{4 \times 3 - 1 \times (-2)}\begin{bmatrix} 4 & -2 \\ 1 & 3 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 4 & -2 \\ 1 & 3 \end{bmatrix}$$

Note:

$$\mathbf{A}\mathbf{A}^{-1} = \begin{bmatrix} 3 & 2 \\ -1 & 4 \end{bmatrix} \cdot \frac{1}{14}\begin{bmatrix} 4 & -2 \\ 1 & 3 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 14 & 0 \\ 0 & 14 \end{bmatrix} = \mathbf{I}$$

No inverse if |**A**| = ad – bc = 0, e.g., $A = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$

Properties of matrix inverse

1. A^{-1} exists & is unique iff (these are all equivalent)
   - (a) $|A| \neq 0$
   - (b) A is non-singular
   - (c) All rows (cols) of A are linearly independent

2. $I^{-1} = I$ since $I \cdot I = I$

   $(A^{-1})^{-1} = A$ since $(A^{-1})(A^{-1})^{-1} = I$
   $= (A^{-1})(A) = I$

   $(A')^{-1} = (A^{-1})'$

Properties of matrix inverse

3. $(AB)^{-1} = B^{-1}A^{-1}$ since $(AB)(B^{-1}A^{-1})$
   $= ABB^{-1}A^{-1} = AA^{-1} = I$

4. If $D = \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}$
   & all $d_i \neq 0$ then $D^{-1} = \begin{pmatrix} 1/d_1 & & & 0 \\ & 1/d_2 & & \\ & & \ddots & \\ 0 & & & 1/d_n \end{pmatrix}$

In general, to show or verify that a matrix **K** is the inverse of matrix **L**, show that **K L = L K = I**

Determinants

For any square matrix, A

$$\det(A) \equiv |A| = \text{a scalar function of } a_{ij}$$

2x2

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$-$ $+$

each term has 1 element from each row & col

e.g. $\begin{vmatrix} 3 & 7 \\ 2 & 8 \end{vmatrix} = 3\cdot 8 - 7\cdot 2 = 10$

Determinants

3x3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

e.g.

$1\cdot 2\cdot 5 \quad 4\cdot 3\cdot 1 \quad 2\cdot 4\cdot 3$

$$\begin{vmatrix} 1 & 3 & 5 \\ 4 & 2 & 3 \\ 1 & 4 & 2 \end{vmatrix} \begin{matrix} 1 & 3 \\ 4 & 2 \\ 1 & 4 \end{matrix} = 4 + 9 + 80 - 10 - 12 - 24 = 93 - 46 = 47$$

$1\cdot 2\cdot 2 \quad 3\cdot 3\cdot 1 \quad 5\cdot 4\cdot 4$

$+ \quad + \quad +$

Determinants: cofactors

- General method: expand by cofactors of a given row or column
  - Minor of a_{ij} : M_{ij} = determinant of submatrix removing the ith row and jth column of **A**.
  - Cofactor of a_{ij} : $C_{ij} = (-1)^{(i+j)} M_{ij}$
  - For row I : $\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots a_{in}C_{in}$
  - For col j: $\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$

Determinants: cofactors

e.g., $A = \begin{bmatrix} 4 & 2 & 1 \\ -2 & -6 & 3 \\ -7 & 5 & 0 \end{bmatrix}$

Minor of a_{12}: $M_{12} = \begin{vmatrix} -2 & 3 \\ -7 & 0 \end{vmatrix}$

Cofactor of a_{12}: $C_{12} = (-1)^{1+2}\begin{vmatrix} -2 & 3 \\ -7 & 0 \end{vmatrix}$

Expand by row 1:

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

$$\det(A) = (4)(+1)\begin{vmatrix} -6 & 3 \\ 5 & 0 \end{vmatrix} + (2)(-1)\begin{vmatrix} -2 & 3 \\ -7 & 0 \end{vmatrix} + (1)(+1)\begin{vmatrix} -2 & -6 \\ -7 & 5 \end{vmatrix}$$

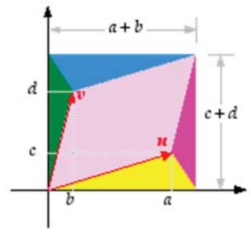
$$= 4(-15) - 2(21) + (1)(-52)$$

$$= -154$$

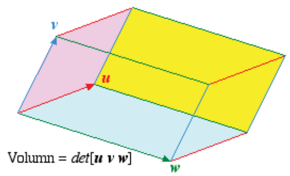
Determinants: geometry

2D: det() = area of parallelogram

$$\det[\mathbf{u}\quad \mathbf{v}] = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$



3D: det() = volume



(What happens if **u**, **v**, **w** are linearly dependent?)

⋮

nD: det() = hyper-volume

Correlation matrices:

$$\det\begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} = 1 - r_{12}^2$$

In general: $0 \le \det(\mathbf{R}_{p\times p}) \le 1$

Singular   Uncorrelated, **R=I**

Summary

- Matrices & vectors: shorthand notation
  - Matrix: 2-way table; vector: 1-way collection of #s
  - Algebra:
    - Addition, subtraction: like ordinary arithmetic
    - Multiplication: **a’ x** = linear combination; **A x** = set of them
  - Use: represent a linear model: $\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
- Inverse: Use to solve linear equations
  - $\mathbf{A}\,\mathbf{x} = \mathbf{b} \rightarrow \mathbf{x} = \mathbf{A}^{-1}\,\mathbf{b}$
- Determinant: “size” of a square matrix
  - | **A** | = 0 → “singular,” no inverse, can’t solve
  - Rank = # linearly independent rows, cols, equations