Linear Algebra

Motivation

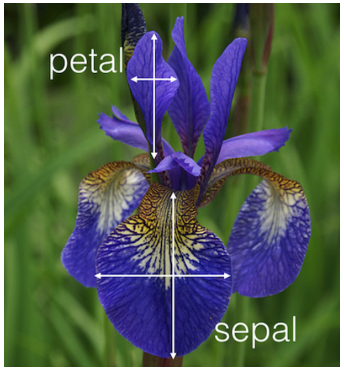
- Just about all real world signals can be represented as vectors and most ordered data sets are organized in tables or “matrices”
- Linear Algebra provides the mathematical background for manipulating and extracting information from such data

Data can be represented by $n \times d$ matrix:

$$\mathbf{D} = \begin{bmatrix} 5.9 & 3.0 & 4.2 & 1.5 \\ 6.9 & 3.1 & 4.9 & 1.5 \\ \vdots & \vdots & \vdots & \vdots \\ 5.1 & 3.4 & 1.5 & 0.2 \end{bmatrix}$$

	Sepal length X_1	Sepal width X_2	Petal length X_3	Petal width X_4
$\mathbf{x}_1$	5.9	3.0	4.2	1.5
$\mathbf{x}_2$	6.9	3.1	4.9	1.5
$\mathbf{x}_3$	6.6	2.9	4.6	1.3
$\mathbf{x}_4$	4.6	3.2	1.4	0.2
$\mathbf{x}_5$	6.0	2.2	4.0	1.0
$\mathbf{x}_6$	4.7	3.2	1.3	0.2
$\mathbf{x}_7$	6.5	3.0	5.8	2.2
$\mathbf{x}_8$	5.8	2.7	5.1	1.9
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{x}_{149}$	7.7	3.8	6.7	2.2
$\mathbf{x}_{150}$	5.1	3.4	1.5	0.2

Extract from Iris data set.



Linear Algebra in the Real World

- Ranking web pages in order of importance (PageRank)
  - Maybe solved as the problem of finding the eigenvector of the page score matrix.

    [The 25,000,000,000 Eigenvector - The Linear Algebra Behind Google]
- Movie recommendation
  - Use Singular Value Decomposition (SVD) to break down User-Item (Movie) matrix into user-feature and item-feature matrices, keeping only the top k-ranks to surface the best matches.

    [Simon Funk, Netflix Prize - http://sifter.org/~simon/journal/20061211.html]
- Topic modeling
  - Extensive use of SVD and matrix factorization can be found in Natural Language Processing, specifically topic modeling and semantic analysis.

Basic Notation

Matrix is an array of number with n rows d columns:

$$\mathbf{D} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1\mathrm{d}} \\ x_{21} & x_{22} & \cdots & x_{2\mathrm{d}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} = [x_{ij}]$$

$$i = 1,2,\cdots n$$
$$j = 1,2,\cdots d.$$

Matrix *transpose:* a matrix with rows and columns interchanged.

$$\mathbf{D}^T = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1\mathrm{n}} \\ x_{21} & x_{22} & \cdots & x_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & \cdots & x_{dn} \end{bmatrix} = [x_{ji}]$$

Basic Notation

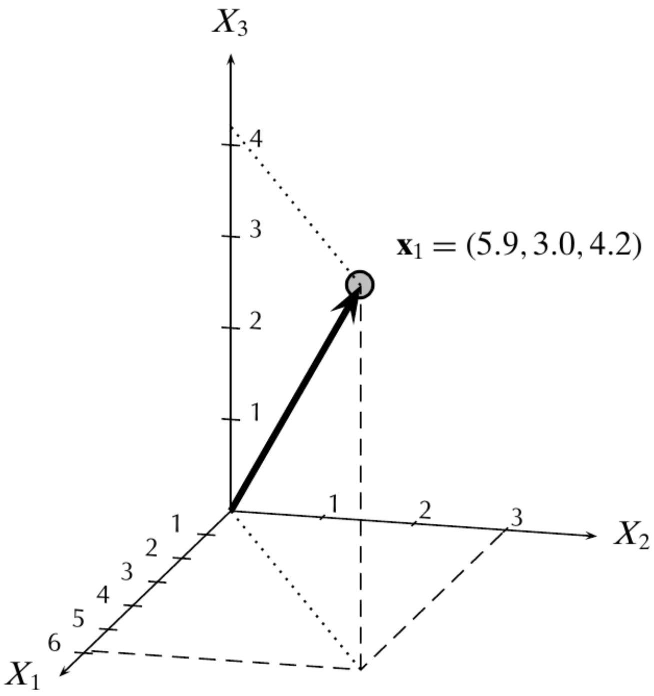
Vector is a matrix with d rows 1 column, also called a column vector:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

Geometrically a vector can be thought of as a point in d-dimensional space

Vector transpose or row vector:

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix}$$



3 dimensional vector

Basic Notation

We can denote the ith row of $\mathbf{D}$ by $\mathbf{x}_i^T$

$$\mathbf{D} = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix}$$

$\mathbf{x}_i^T$ are d dimensional row vectors.

We can denote the jth column of $\mathbf{D}$ by $\mathbf{y}_j$

$$\mathbf{D} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_d \end{bmatrix}$$

$\mathbf{y}_j$ are n dimensional column vectors.

Addition and Scalar Multiplication

The sum of two matrices $\mathbf{A}$ and $\mathbf{B}$ *of the same size* is written as

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

where each element of $\mathbf{C}$ is $c_{ij} = a_{ij} + b_{ij}$

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

$$\mathbf{C} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -4 & 6 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 5 & -1 \\ 3 & 1 \end{bmatrix}$$

Addition and Scalar Multiplication

The product of any matrix $\mathbf{A}$ any scalar c (number c) is written $c\mathbf{A}$ obtained by multiplying each entry of $\mathbf{A}$ by c

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad c\,\mathbf{A} = \begin{bmatrix} c\,a_{11} & c\,a_{12} \\ c\,a_{21} & c\,a_{22} \end{bmatrix}$$

Rules for Matrix Addition and Scalar Multiplication

$$\text{(a)} \quad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$
$$\text{(b)} \quad (\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$
$$\text{(c)} \quad c(\mathbf{A} + \mathbf{B}) = c\,\mathbf{A} + c\,\mathbf{B}$$

Matrix Multiplication

The product of an $m \times n$ matrix $\mathbf{A} = [a_{ij}]$ and an $n \times p$ matrix $\mathbf{B} = [b_{ij}]$ is a $m \times p$ matrix:

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} = [c_{ij}]$$

with entries:

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj} \qquad \begin{array}{l} i = 1, 2, \cdots m \\ j = 1, 2, \cdots p. \end{array}$$

Note that in order for the matrix product to exist, the number of columns in $\mathbf{A}$ must equal the number of rows in $\mathbf{B}$.

$$\begin{array}{ccccc} \mathbf{A} & \mathbf{B} & = & \mathbf{C} \\ [m \times n] & [n \times p] & = & [m \times p] \end{array}$$

$$m = 4 \left\{ \overbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}}^{n\,=\,3} \right. \overbrace{\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}}^{p\,=\,2} = \left. \overbrace{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix}}^{p\,=\,2} \right\} m$$

$$c_{21} = a_{21} b_{11} + a_{22} b_{21} + a_{23} b_{31}$$

Vector-Vector Product

Given two vectors x,y we can define two products:

inner product or *dot product*

(both vector must have same dimensions)

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{bmatrix} = \sum_{i=1}^{d} x_i y_i$$

Note: $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$

outer product

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \cdots & x_m y_n \end{bmatrix}$$

Matrix-Vector Product

If we write matrix $\mathbf{A}$ in row form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & x_{12} & \cdots & a_{1\mathrm{d}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{d}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nd} \end{bmatrix}$$

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix}$$

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{x} \\ \mathbf{a}_2^T \mathbf{x} \\ \vdots \\ \mathbf{a}_n^T \mathbf{x} \end{bmatrix}$$

Alternatively, if we write $\mathbf{A}$ in column form:

$$\mathbf{y} = \mathbf{A}\mathbf{x} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_d\boldsymbol{\alpha}_d$$

$\mathbf{y}$ can be written as a linear combination of column vectors of $\mathbf{A}$.

Matrix-Matrix Multiplication

Product in terms of row and column vectors:

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1^T\mathbf{b}_1 & \mathbf{a}_1^T\mathbf{b}_2 & \cdots & \mathbf{a}_1^T\mathbf{b}_n \\ \mathbf{a}_2^T\mathbf{b}_1 & \mathbf{a}_2^T\mathbf{b}_2 & \cdots & \mathbf{a}_2^T\mathbf{b}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m^T\mathbf{b}_1 & \mathbf{a}_m^T\mathbf{b}_2 & \cdots & \mathbf{a}_m^T\mathbf{b}_n \end{bmatrix}$$

Matrix-matrix multiplication as a set of matrix-vector products

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} = \mathbf{A}\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}\,\mathbf{b}_1 & \mathbf{A}\,\mathbf{b}_2 & \cdots & \mathbf{A}\,\mathbf{b}_n \end{bmatrix}$$

Some properties of matrix multiplication:

(a) $\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}$

(b) $(\mathbf{A}\,\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\,\mathbf{C}) = \mathbf{A}\,\mathbf{B}\,\mathbf{C}$

(c) $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{A}\,\mathbf{C} + \mathbf{B}\,\mathbf{C}$

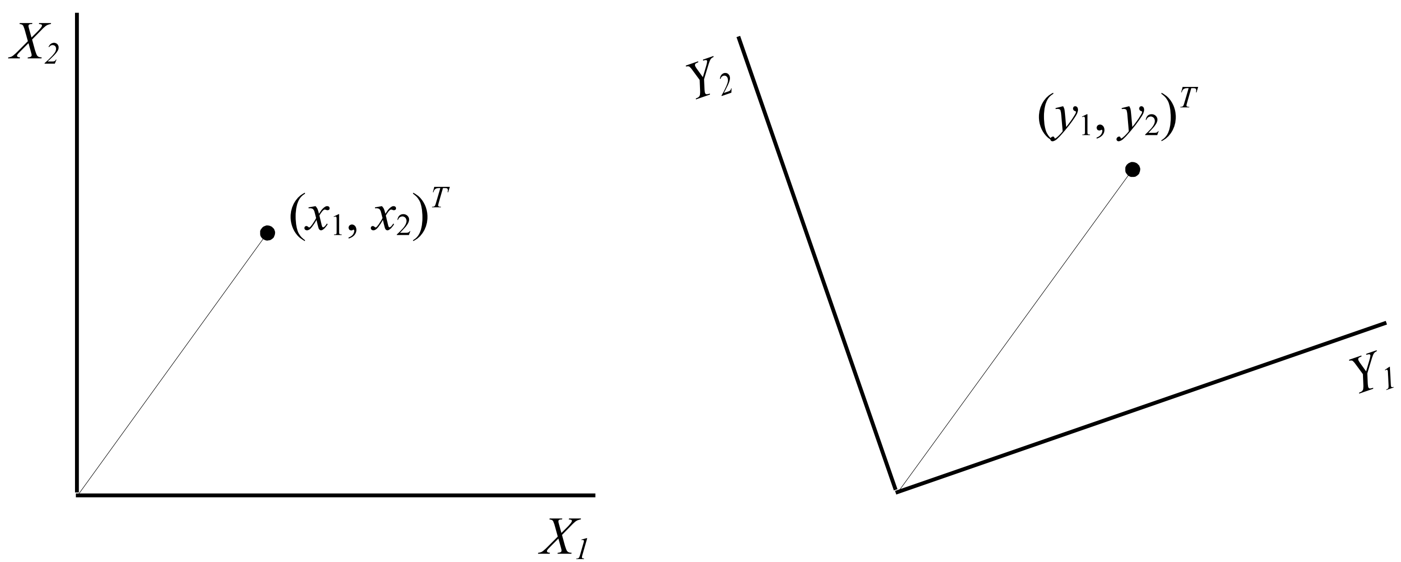
Why Matrix Multiplication (Transformation)

We can motivate the "unnatural" matrix multiplication by its use in linear transformations:

$$y_1 = a_{11}x_1 + a_{12}x_2$$

$$y_2 = a_{21}x_1 + a_{22}x_2$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{A}\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$



Why Matrix Multiplication (Transformation)

Suppose that X_1X_2-system is related to W_1W_2-system:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{B}\mathbf{w} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} b_{11}w_1 + b_{12}w_2 \\ b_{21}w_1 + b_{22}w_2 \end{bmatrix}$$

Then the Y_1Y_2-system is related to the W_1W_2-system indirectly via the X_1X_2-system:

$$\begin{aligned} y_1 &= a_{11}(b_{11}w_1 + b_{12}w_2) + a_{12}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{11}b_{11} + a_{12}b_{21})w_1 + (a_{11}b_{12} + a_{12}b_{22})w_2 \\ y_2 &= a_{21}(b_{11}w_1 + b_{12}w_2) + a_{22}(b_{21}w_1 + b_{22}w_2) \\ &= (a_{21}b_{11} + a_{22}b_{21})w_1 + (a_{21}b_{12} + a_{22}b_{22})w_2 \end{aligned}$$

Or we can use matrix multiplication!

$$\mathbf{y} = \mathbf{A}\,\mathbf{B}\,\mathbf{w}$$

Matrix Transpose

Matrix *transpose:* a matrix with rows and columns interchanged.

$$\mathbf{D}^T = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1\mathrm{n}} \\ x_{21} & x_{22} & \cdots & x_{2\mathrm{d}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & \cdots & x_{dn} \end{bmatrix} = \left[x_{ji}\right]$$

Rules for transposition:

(a) $\left(\mathbf{A}^T\right)^T = \mathbf{A}$

(b) $\left(\mathbf{A}+\mathbf{B}\right)^T = \mathbf{A}^T + \mathbf{B}^T$

(c) $\left(\mathbf{A}\,\mathbf{B}\right)^T = \mathbf{B}^T\,\mathbf{A}^T$

Special Matrices

Symmetric Matrix is a square matrix whose transpose equals the matrix itself:

$$\mathbf{A}^T = \mathbf{A}$$

Note: $\mathbf{D}^T\mathbf{D}$ and $\mathbf{D}\mathbf{D}^T$ are always symmetric!

Diagonal Matrix is a square matrix that can have non-zero entries only on the main diagonal. *e.g.*,

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

identity matrix

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

Special Matrix

Triangular matrices are square matrices that can have nonzero entries only on and *above* or *below* the main diagonal.

$$\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 6 \end{bmatrix}, \qquad \begin{bmatrix} 2 & 0 & 0 \\ 8 & -1 & 0 \\ 7 & 6 & 8 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 & 0 \\ 9 & -3 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 9 & 3 & 6 \end{bmatrix}$$

Upper triangular

Lower triangular

Norm

The *Euclidean norm* or length of a vector is defined as:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

Unit vector: $\|\mathbf{u}\| = 1$

$$\mathbf{u} = \frac{1}{\|\mathbf{x}\|}\mathbf{x}$$

$\mathbf{u}$ gives the direction of $\mathbf{x}$.

The Euclidean norm is a special case of a general class of norms, know as L_p-norm:

$$\|\mathbf{x}\|_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_d|^p \right)^{1/p}$$

for any $p \neq 0$. Thus, the Euclidean norm corresponds to the case when $p=2$.

Distance and Angle

The *Euclidean distance* between two vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

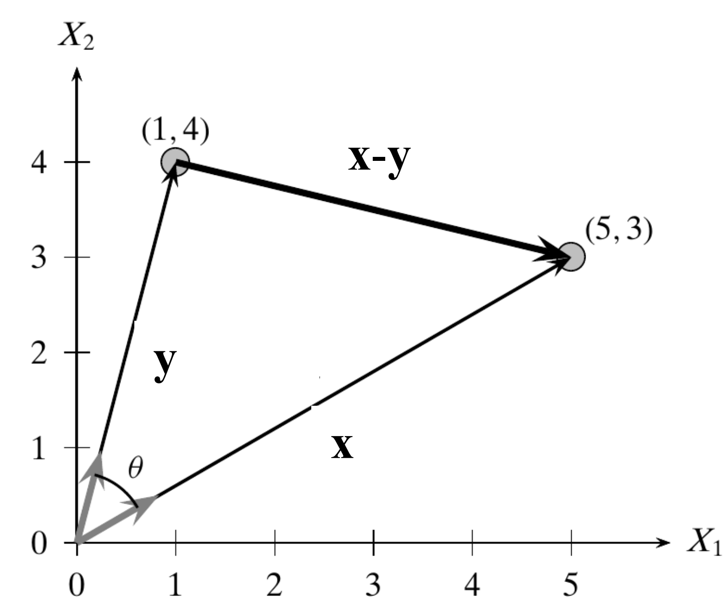
$$\|\mathbf{x}-\mathbf{y}\| = \sqrt{(\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})}$$

Cosine similarity: the smallest angle between vectors $\mathbf{x}$ and $\mathbf{y}$:

$$\cos\theta = \frac{\mathbf{x}^T\mathbf{y}}{\|\mathbf{x}\| \; \|\mathbf{y}\|} = \left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)^T\left(\frac{\mathbf{y}}{\|\mathbf{y}\|}\right)$$

→ dot produce of two unit vectors!

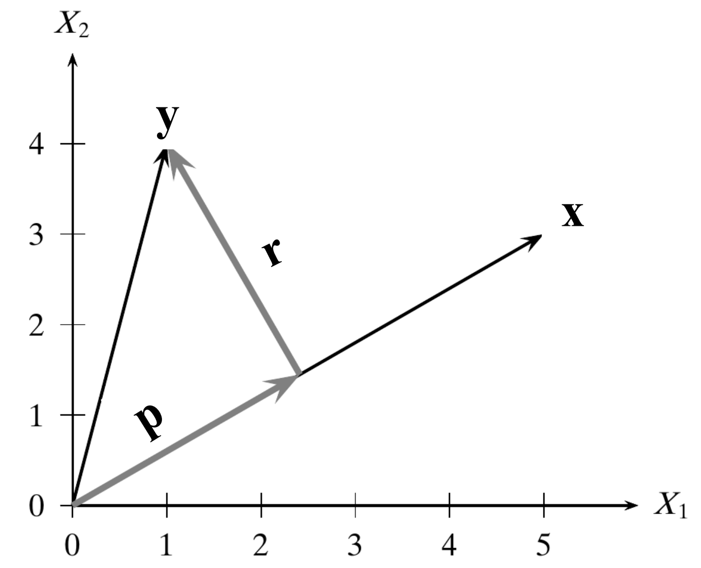
Two vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* if $\mathbf{x}^T\mathbf{y} = 0$, which in turn implies $\cos\theta = 0$ or the angle between them is 90°.
In this case, we say that they have no similarity.



Orthogonal Projection

Consider two d-dimensional vectors $\mathbf{x}$, $\mathbf{y}$
Orthogonal decomposition of $\mathbf{y}$ w.r.t $\mathbf{x}$:

$$\mathbf{y} = \mathbf{p} + \mathbf{r}$$



$\mathbf{r}$ gives the perpendicular distance between $\mathbf{y}$ and $\mathbf{x}$.

$\mathbf{p}$ is the orthogonal projection of $\mathbf{y}$ on $\mathbf{x}$.

$$\mathbf{p} = c\,\mathbf{x} = \left(\frac{\mathbf{x}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x}}\right)\mathbf{x}$$

Mean and Variance of Data Matrix

The mean of the data matrix D is a vector obtained as the average of all the points:

$$mean(\mathbf{D}) = \boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i$$
$$= (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_d)$$

$$\mathbf{D} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix}$$

(*i.e.*, mean of every column)

Total variance of the data matrix $\mathbf{D}$ is the average squared distance of each point from the mean:

$$var(\mathbf{D}) = \frac{1}{n}\sum_{i=1}^{n}\|\mathbf{x}_i - \boldsymbol{\mu}\|^2 = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})^T(\mathbf{x}_i - \boldsymbol{\mu})$$

$$= \frac{1}{n}\sum_{i=1}^{n}\|\mathbf{x}_i\|^2 - \|\boldsymbol{\mu}\|^2$$

Mean-Centered Data

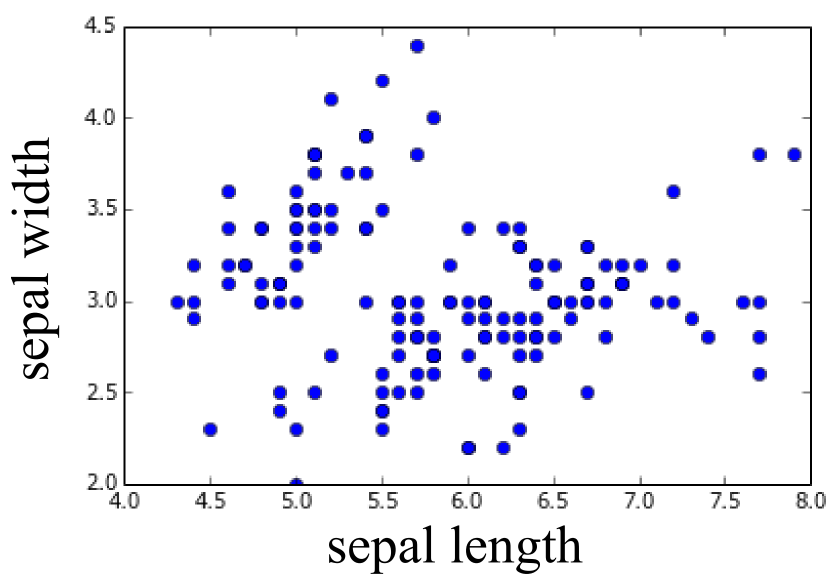
Often we need to center the data matrix by making the mean coincide with the origin of the data space.

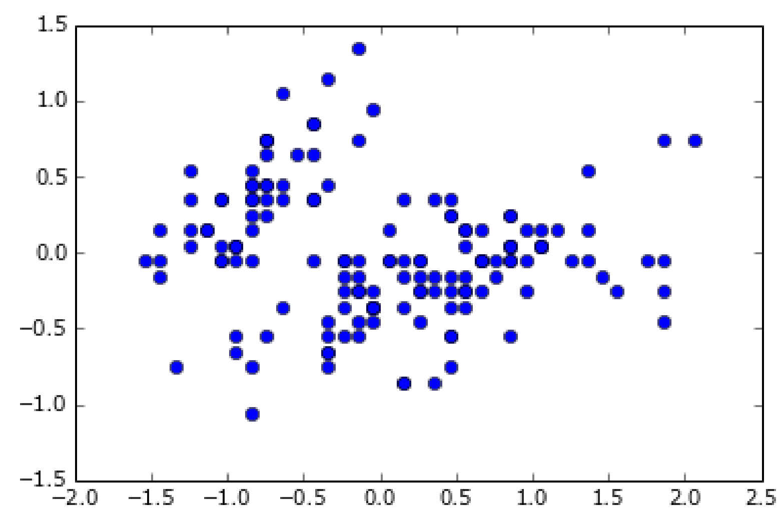
The *centered data matrix* is obtained by subtracting the mean from all the points.

$$\mathbf{Z} = \mathbf{D} - \mathbf{1}\boldsymbol{\mu}^T = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}^T \\ \boldsymbol{\mu}^T \\ \vdots \\ \boldsymbol{\mu}^T \end{bmatrix} \qquad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

n-dim vector

Note: $mean(\mathbf{Z}) = 0$





Intermission!

System of Linear Equations

A system of linear equations, or linear system is a collection of linear equations involving the same set of variables:

$$\begin{aligned} 2x + 3y &= 6 \\ 4x + 5y &= 15 \end{aligned}$$

The general form of a linear system of m equations in n variables is

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{gathered}$$

Linear Equations as Matrices

The linear system can be compactly represented in matrix form as:

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}$$

$$\begin{bmatrix} a_{11} & x_{12} & \cdots & a_{1\mathrm{n}} \\ a_{21} & a_{22} & \cdots & a_{2\mathrm{n}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Solving System of Linear Equations

Consider a simple system:

$$
\begin{aligned}
L_1 &: \quad x + y + z = 6 \\
L_2 &: \quad 2x - y + 4z = 12 \\
L_3 &: \quad 3x + y - 2z = -1
\end{aligned}
$$

Augmented matrix:

$$
\left[\begin{array}{ccc|c}
1 & 1 & 1 & 6 \\
2 & -1 & 4 & 12 \\
3 & 1 & -2 & -1
\end{array}\right]
$$

We can define three *elementary row operations*, which when applied to a linear system do not change the solution of set of the linear system:

- Swapping two rows
- Scalar (non-zero) multiplication of a single row
- Adding a scalar multiple of one row to another row

Solving System of Linear Equations

Step 1 : Express as Augmented Matrix

$$\begin{array}{c} L_1 \\ L_2 \\ L_3 \end{array} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 2 & -1 & 4 & 12 \\ 3 & 1 & -2 & -1 \end{array}\right]$$

Step 2 : Eliminate x from L_2 and L_3

$$\begin{array}{c} L_1 \\ L_2 - 2L_1 \to L_2 \\ L_3 - 3L_1 \to L_3 \end{array} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & -3 & 2 & 0 \\ 0 & -2 & -5 & -19 \end{array}\right]$$

Step 3 : Eliminate y from L_3

$$\begin{array}{c} L_1 \\ L_2 \\ L_3 - \frac{2}{3}L_2 \to L_3 \end{array} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & -\frac{19}{3} & -19 \end{array}\right]$$

$$x + y + z = 1; \quad -3y + 2z = 0; \quad -\frac{19}{3}z = -19$$

Solving System of Linear Equations

At this point, the last row is

$$\frac{-19}{3}z=-19 \quad \Rightarrow \quad z=3$$

Step 4 : Back substitute z into L_2 to find y

$$L_2: \quad -3y+2z=0 \quad \Rightarrow \quad -3\mathrm{y}+6=0 \quad \Rightarrow \quad y=2$$

Step 5 : Back substitute y and z into L_1 to find x

$$L_1: \quad x+y+z=6 \quad \Rightarrow \quad x+2+3=6 \quad \Rightarrow \quad x=1$$

This procedure is called *Gauss Elimination Method*

Determinants

Determinant of a 2×2 matrix

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Determinant of 3×3 matrix

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

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

Determinants

Determinant of a $n \times n$ matrix

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(\mathbf{A}_{1j})$$

$\mathbf{A}_{1j}$, called minor, is a sub-matrix of $\mathbf{A}$.
Example:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\mathbf{A}_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

$$\mathbf{A}_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

Some properties:

(a) $\det(\mathbf{A}^T) = \det(\mathbf{A})$

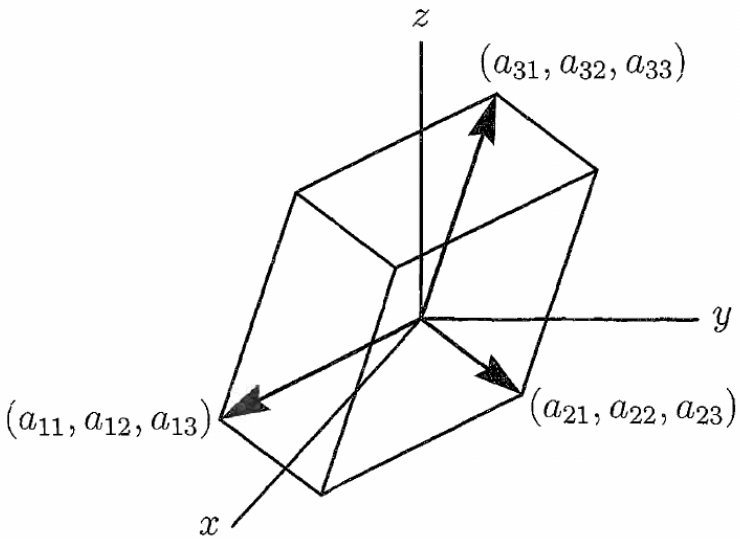
(b) $\det(\mathbf{A}\,\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$

Determinants

Geometrical interpretation:

The determinant of an $n \times n$ matrix is the volume of an n-dimensional *parallelotope* whose sides are given by rows of the matrix!

The determinant of a 3×3 matrix $\mathbf{A}$ gives the volume of *parallelepiped* with sides defined by rows of $\mathbf{A}$.



Determinant

Determinant of a triangular matrix is the product of the diagonal elements!
For example:

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} \qquad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

For an $n \times n$ triangular matrix:

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33}\cdots a_{nn}$$

We can use elementary row operations to convert a square matrix into a triangular matrix.

Matrix Inverse

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

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

If $\mathbf{A}$ has an inverse, then $\mathbf{A}$ is called a *nonsingular matrix*.
If $\mathbf{A}$ has no inverse, then $\mathbf{A}$ is called a *singular matrix*.

If $\mathbf{A}$ is nonsingular, then $\det(\mathbf{A}) \neq 0$

The Matrix Eigenvalue Problem

Consider multiplying nonzero vectors by a given square matrix, such as

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 \\ 27 \end{bmatrix}$$

Compare to this:

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix} = 10\begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

The Matrix Eigenvalue Problem

Consider multiplying nonzero vectors by a given square matrix, such as

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 \\ 27 \end{bmatrix}$$

Compare to this:

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix} = 10\begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Eigenvalue equation:

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}$$

Here $\mathbf{A}$ is a *square* matrix, $\mathbf{x}$ is a row vector, and λ is a scaler.

The problem of systematically finding such λ's and nonzero vectors for a given square matrix is called the matrix eigenvalue problem or, more commonly, the eigenvalue problem.

The Matrix Eigenvalue Problem

Eigenvalue equation:

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}$$

λ is called the eigenvalue of $\mathbf{A}$
$\mathbf{x}$ is called the eigenvector of $\mathbf{A}$
corresponding to eigenvalue λ.

Remark: $\mathbf{x} = 0$ is always a solution, but this is of no interest.

How to Find Eigenvalues and Eigenvectors

Lets consider a simple example:

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

Eigenvalues must be determined first:

$$\mathbf{A}\,\mathbf{x} = \lambda\,\mathbf{x}$$

$$\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

In component form:

$$\begin{aligned} -5\,x_1 + 2\,x_2 &= \lambda\,x_1 \\ 2\,x_1 - 2\,x_2 &= \lambda\,x_2 \end{aligned}$$

$$\begin{aligned} (-5-\lambda)\,x_1 \; + \; 2\,x_2 &= 0 \\ 2\,x_1 \; + \; (-2-\lambda)\,x_2 &= 0 \end{aligned}$$

How to Find Eigenvalues and Eigenvectors

Back to matrix form:

$$\begin{bmatrix} (-5-\lambda) & 2 \\ 2 & (-2-\lambda) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(\mathbf{A} - \lambda\,\mathbf{I})\,\mathbf{x} \;=\; \mathbf{0} \qquad\qquad \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This is a *homogeneous* equation.

By *Cramer's rule*, the solution exists if and only if the determinant of coefficient matrix is zero.

$$\det(\mathbf{A} - \lambda\,\mathbf{I}) \;=\; 0$$ → *Characteristic equation*

$$\begin{vmatrix} (-5-\lambda) & 2 \\ 2 & (-2-\lambda) \end{vmatrix} \;=\; 0$$

How to Find Eigenvalues and Eigenvectors

$$\begin{vmatrix} (-5-\lambda) & 2 \\ 2 & (-2-\lambda) \end{vmatrix} = 0$$

$$(-5-\lambda)(-2-\lambda) - 4 = 0$$

$$\lambda^2 + 7\lambda + 6 = 0$$

→ quadratic equation.

The solutions of this equation are: $\lambda_1 = -1, \quad \lambda_2 = -6$

These are Eigenvalues of $\mathbf{A}$.

How to Find Eigenvalues and Eigenvectors

Starting from the linear equations:

$$\begin{aligned}(-5-\lambda)x_1 \; + \; 2x_2 &= 0 \\ 2x_1 \; + \; (-2-\lambda)x_2 &= 0\end{aligned}$$

Eigenvector of $\mathbf{A}$ corresponding to $\lambda = \lambda_1 = -1$

$$\begin{aligned}-4x_1 + 2x_2 &= 0 \\ 2x_1 - x_2 &= 0\end{aligned}$$

$$\Rightarrow \quad x_2 = 2x_1$$

This determines an eigenvector up to a scalar multiple.

Choose $x_1 = 1$

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Check:

$$\mathbf{A}\,\mathbf{x}_1 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

How to Find Eigenvalues and Eigenvectors

Starting from the linear equations:

$$\begin{aligned}(-5-\lambda)x_1 &+ 2x_2 = 0\\ 2x_1 &+ (-2-\lambda)x_2 = 0\end{aligned}$$

Eigenvector of $\mathbf{A}$ corresponding to $\lambda = \lambda_2 = -6$

$$\begin{aligned}x_1 + 2x_2 &= 0\\ 2x_1 - 4x_2 &= 0\end{aligned}$$

$$\Rightarrow \quad x_1 = -2x_2$$

Choose $x_2 = 1$

$$\mathbf{x}_2 = \begin{bmatrix}-2\\ 1\end{bmatrix}$$

Check:

$$\mathbf{A}\,\mathbf{x}_2 = \begin{bmatrix}-5 & 2\\ 2 & -2\end{bmatrix}\begin{bmatrix}-2\\ 1\end{bmatrix} = \begin{bmatrix}12\\ -6\end{bmatrix} = (-6)\begin{bmatrix}-2\\ 1\end{bmatrix}$$

How to Find Eigenvalues and Eigenvectors

The general case:

$$\begin{bmatrix} a_{11} & x_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \longrightarrow$$

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= \lambda x_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= \lambda x_2 \\ \cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= \lambda x_n. \end{aligned}$$

$$\begin{aligned} (a_{11}-\lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n &= 0. \end{aligned}$$

How to Find Eigenvalues and Eigenvectors

In matrix notation:

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

This has a solution if and only if:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \qquad \rightarrow \textit{Characteristic equation}$$

This gives rise to a polynomial of degree n in λ.

The eigenvalues of a square matrix $\mathbf{A}$ are the roots of the characteristic equation of $\mathbf{A}$.

Hence an $n \times n$ matrix has at least one eigenvalue and at most n numerically different eigenvalues.

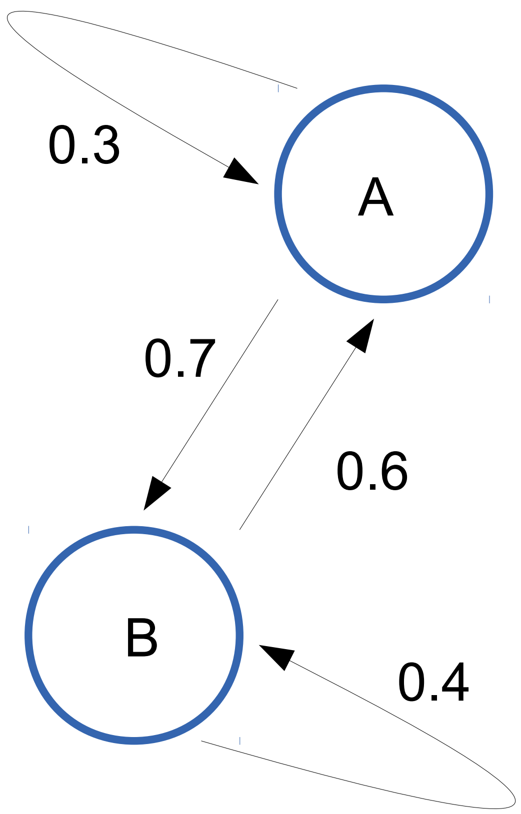
The eigenvalues must be determined first. Once these are known, corresponding eigenvectors are obtained from the system of n linear equations.

Markov Process

Markov Processes, also called *Markov Chains* are described as a series of *states* which transition from one to another, and have a given probability for each transition.

Transition Matrix

$$\begin{array}{cc} & \text{Current State} \\ & \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \text{Future State} \quad \begin{array}{c} \text{A} \\ \text{B} \end{array} & \begin{bmatrix} 0.3 & 0.6 \\ 0.7 & 0.4 \end{bmatrix} \end{array}$$



$$\mathbf{x}_{i+1} = \mathbf{A}\,\mathbf{x}_i$$

$$\begin{bmatrix} 0.3 & 0.6 \\ 0.7 & 0.4 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}$$

Stationary State:

$$\mathbf{A}\,\mathbf{x} = 1\,\mathbf{x}$$