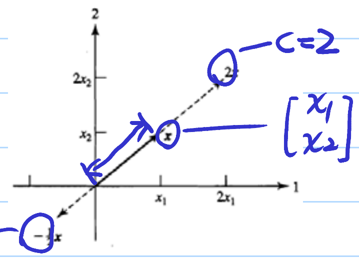
Matrix Algebra

- vector (向量)
  - An array $\mathbf{x}$ of n real numbers $x_1, x_2, \ldots, x_n$ is called a *vector*, and it is written as

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



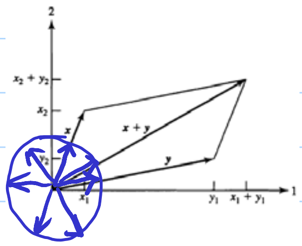
two basic operations

- scalar multiplication

  $c\mathbf{x}$ is the vector obtained by multiplying each element of $\mathbf{x}$ by c. $c\mathbf{x} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$

- addition

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$



- a vector has both length and direction
  - $L_{\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x'x}$
  - $L_{c\mathbf{x}} = |c| L_{\mathbf{x}}$
  - multiplication by c does not change the direction of vector
  - unit vectors on the direction of $\mathbf{x}$: $L_{\mathbf{x}}^{-1}\mathbf{x}$ length of $L_x^{-1}x = 1$.

- vector space

  Definition 2A.4. The space of all real m-tuples, with scalar multiplication and vector addition as just defined, is called a *vector space*.

  Definition 2A.5. The vector $\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k$ is a *linear combination* of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. The set of all linear combinations of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, is called their *linear span*. $\dim(\mathrm{span}(x_1, \cdots, x_k)) \le k$

  $x_3 = a_1x_1 + a_2x_2$

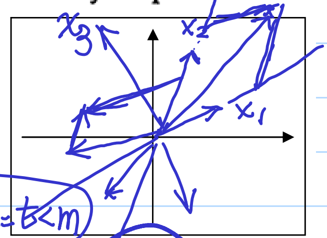
- linearly dependent and independent

  $x_3 - a_1x_1 - a_2x_2 = 0$

  Definition 2A.6. A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is said to be *linearly dependent* if there exist k numbers $(a_1, a_2, \ldots, a_k)$, not all zero, such that

  matrix explanation $= \underbrace{[x_1 | x_2 | \cdots | x_k]}_{m\times k}$

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k = \mathbf{0}$$



  Otherwise the set of vectors is said to be *linearly independent*.

- basis of vector space

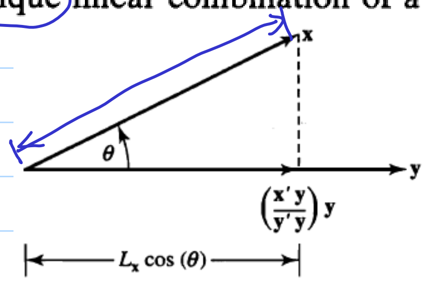
  m vectors, linearly dependent, dim = $k < m$

  Definition 2A.7. Any set of m linearly independent vectors is called a *basis* for the vector space of all m-tuples of real numbers. 基底

  Result 2A.1. Every vector can be expressed as a unique linear combination of a fixed basis.

- inner product (內積)

$$\mathbf{x}'\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$
$$= L_{\mathbf{x}}L_{\mathbf{y}} \cos(\theta)$$



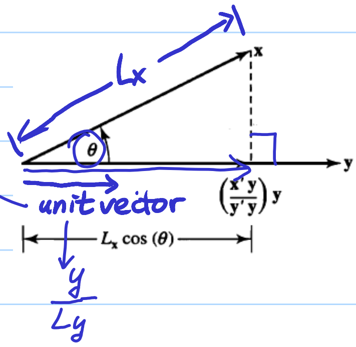
  - length: $L_{\mathbf{x}} = \text{length of } \mathbf{x} = \sqrt{\mathbf{x}'\mathbf{x}}$

- angle between 2 vectors: $\cos(\theta) = \dfrac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}}L_{\mathbf{y}}} = \dfrac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}$ c.f. → sample correlation
- projection of $\mathbf{x}$ on $\mathbf{y}$:

$$\text{Projection of } \mathbf{x} \text{ on } \mathbf{y} = \frac{(\mathbf{x}'\mathbf{y})}{\mathbf{y}'\mathbf{y}}\mathbf{y} = \boxed{\frac{(\mathbf{x}'\mathbf{y})}{L_{\mathbf{y}}}}\boxed{\frac{1}{L_{\mathbf{y}}}\mathbf{y}}$$

unit vector

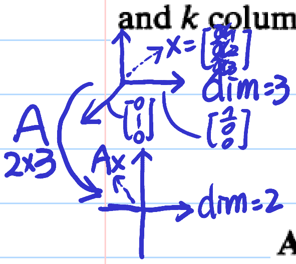
$$\text{Length of projection} = \frac{|\mathbf{x}'\mathbf{y}|}{L_{\mathbf{y}}} = L_{\mathbf{x}}\left|\frac{\mathbf{x}'\mathbf{y}}{L_{\mathbf{x}}L_{\mathbf{y}}}\right| = L_{\mathbf{x}}|\cos(\theta)|$$



- matrix

➢ **Definition 2A.13.** An $\boxed{m \times k}$ *matrix*, generally denoted by a boldface uppercase letter such as $\mathbf{A}, \mathbf{R}, \mathbf{\Sigma}$, and so forth, is a rectangular array of elements having m rows and k columns.

← dimension of matrix



a column — a row

$$\underset{(m\times k)}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix} \qquad \underset{(m\times k)}{\mathbf{A}} = \{a_{ij}\}$$

$$\mathbf{A}\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_k\mathbf{a}_k, \text{ where } \mathbf{a}_i \text{ is the } i\text{th column of } \mathbf{A}.$$

➢ transpose

Definition 2A.19. Consider the $m \times k$ matrix $\mathbf{A}$ with arbitrary elements a_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, k$. The *transpose* of the matrix $\mathbf{A}$, denoted by $\mathbf{A}'$, is the $k \times m$ matrix with elements a_{ji}, $j = 1, 2, \ldots, k$, $i = 1, 2, \ldots, m$. That is, the transpose of the matrix $\mathbf{A}$ is obtained from $\mathbf{A}$ by interchanging the rows and columns.

➢ addition

Definition 2A.16 (Matrix addition). Let the matrices $\mathbf{A}$ and $\mathbf{B}$ both be of dimension $m \times k$ with arbitrary elements a_{ij} and b_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, k$, respectively. The sum of the matrices $\mathbf{A}$ and $\mathbf{B}$ is an $m \times k$ matrix $\mathbf{C}$, written $\mathbf{C} = \mathbf{A} + \mathbf{B}$, such that the arbitrary element of $\mathbf{C}$ is given by

$A+B=\{a_{ij}+b_{ij}\}$
$A-B=\{a_{ij}-b_{ij}\}$

$$c_{ij} = a_{ij} + b_{ij} \qquad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, k$$

➢ scalar multiplication

Definition 2A.17 (Scalar multiplication). Let c be an arbitrary scalar and $\underset{(m\times k)}{\mathbf{A}} = \{a_{ij}\}$. Then $\underset{(m\times k)}{c\mathbf{A}} = \underset{(m\times k)}{\mathbf{A}c} = \underset{(m\times k)}{\mathbf{B}} = \{b_{ij}\}$, where $b_{ij} = ca_{ij} = a_{ij}c$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, k$.

➢ matrix multiplication

Definition 2A.23 (Matrix multiplication). The product $\mathbf{AB}$ of an $m \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ and an $n \times k$ matrix $\mathbf{B} = \{b_{ij}\}$ is the $m \times k$ matrix $\mathbf{C}$ whose elements are

$$c_{ij} = \sum_{\ell=1}^{n} a_{i\ell}b_{\ell j} \qquad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, k$$

$$\underset{(n\times4)(4\times p)}{\mathbf{A}\ \mathbf{B}'} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & a_{i4} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ b_{31} & \cdots & b_{3j} & \cdots & b_{3p} \\ b_{41} & \cdots & b_{4j} & \cdots & b_{4p} \end{bmatrix} = \begin{array}{c} \\ \\ \text{Row } i \end{array} \begin{array}{c} \text{Column} \\ j \\ \left[\begin{array}{c} \vdots \\ \cdots (a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + a_{i4}b_{4j}) \cdots \\ \vdots \end{array} \right] \end{array}$$

inner product.

- in general, $\mathbf{AB} \neq \mathbf{BA}$ → e.g. A: m×k, B: k×m; AB = m×m, BA = k×k, m≠k.

- some properties of matrix operations

Result 2A.4. For all matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ (of equal dimension) and scalars c and d, the following hold:

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

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

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

(d) $(c+d)\mathbf{A} = c\mathbf{A}+d\mathbf{A}$

(e) $(\mathbf{A}+\mathbf{B})' = \mathbf{A}'+\mathbf{B}'$ (That is, the transpose of the sum is equal to the sum of the transposes.)

(f) $(cd)\mathbf{A} = c(d\mathbf{A})$

(g) $(c\mathbf{A})' = c\mathbf{A}'$

Result 2A.5. For all matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ (of dimensions such that the indicated products are defined) and a scalar c,

(a) $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B}$

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

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

(d) $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$

(e) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ $\neq A'B'$

More generally, for any $\mathbf{x}_j$ such that $\mathbf{A}\mathbf{x}_j$ is defined,

(f) $\sum_{j=1}^{n} \mathbf{A}\mathbf{x}_j = \mathbf{A}\sum_{j=1}^{n}\mathbf{x}_j$

(g) $\sum_{j=1}^{n} (\mathbf{A}\mathbf{x}_j)(\mathbf{A}\mathbf{x}_j)' = \mathbf{A}\left(\sum_{j=1}^{n}\mathbf{x}_j\mathbf{x}_j'\right)\mathbf{A}'$

$x_j' A'$

$\begin{bmatrix} Ax_1 \\ Ax_k \end{bmatrix}$ $[Ax_1 \ \ Ax_k]$ $\longleftrightarrow$ inner product

dim(span(rows))

dim(span(columns))

- rank

Definition 2A.25. The *row rank* of a matrix is the maximum number of linearly independent rows, considered as vectors (that is, row vectors). The *column rank* of a matrix is the rank of its set of columns, considered as vectors.

Result 2A.6. The row rank and the column rank of a matrix are equal.

rank of a matrix is either the row rank or the column rank

- square matrix: # of rows = # of columns $(m=k)$
  - a square matrix is said to be symmetric if $a_{ij}=a_{ji}$ $(A'=A)$
  - identity matrix $\mathbf{I}$: the square matrix with ones on the diagonal and zero elsewhere

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & & a_{2p} \\ \vdots & & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}$$

$$\underset{(k\times k)(k\times k)}{\mathbf{I}\ \ \mathbf{A}} = \underset{(k\times k)(k\times k)}{\mathbf{A}\ \ \mathbf{I}} = \underset{(k\times k)}{\mathbf{A}} \quad \text{for any } \underset{(k\times k)}{\mathbf{A}}$$

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

- singularity

non singular: columns of A are linearly independent

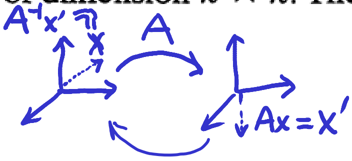
Definition 2A.26. A square matrix $\underset{(k\times k)}{\mathbf{A}}$ is *nonsingular* if $\boxed{\underset{(k\times k)(k\times 1)}{\mathbf{A}\ \ \mathbf{x}}} = \underset{(k\times 1)}{\mathbf{0}}$ implies that $\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{\mathbf{0}}$. If a matrix fails to be nonsingular, it is called *singular*. Equivalently, a *square* matrix is nonsingular if its rank is equal to the number of rows (or columns) it has.

- inverse of a square matrix

Result 2A.7. Let $\mathbf{A}$ be a nonsingular square matrix of dimension $k \times k$. Then there is a unique $k \times k$ matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}$$

where $\mathbf{I}$ is the $k \times k$ identity matrix. exchangable



Definition 2A.27. The $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ is called the *inverse* of $\mathbf{A}$ and is denoted by $\mathbf{A}^{-1}$. In fact, if $\mathbf{BA} = \mathbf{I}$ *or* $\mathbf{AB} = \mathbf{I}$, then $\mathbf{B} = \mathbf{A}^{-1}$, and both products must equal $\mathbf{I}$.

Result 2A.9. For a square matrix $\mathbf{A}$ of dimension $k \times k$, the following are equivalent:

(a) $\underset{(k\times k)(k\times 1)}{\mathbf{A}\ \ \mathbf{x}} = \underset{(k\times 1)}{\mathbf{0}}$ implies $\underset{(k\times 1)}{\mathbf{x}} = \underset{(k\times 1)}{\mathbf{0}}$ ($\mathbf{A}$ is nonsingular).

(b) $|\mathbf{A}| \neq 0$.

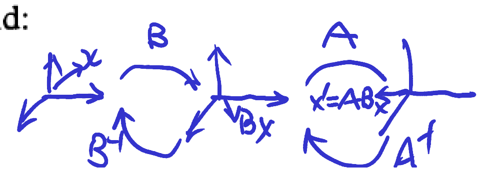
(c) There exists a matrix $\mathbf{A}^{-1}$ such that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \underset{(k\times k)}{\mathbf{I}}$.

Result 2A.10. Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of the same dimension, and let the indicated inverses exist. Then the following hold:

(a) $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$

(b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ $\neq A^{-1}B^{-1}$

ABx



- orthogonal square matrix

Definition 2A.29. A square matrix $\mathbf{A}$ is said to be *orthogonal* if its rows, considered as vectors, are mutually perpendicular and have unit lengths; that is, $\mathbf{AA}' = \mathbf{I}$.

$\hookrightarrow x_i' x_j' = 0$ if $i \neq j$. $x_j \perp x_i$ ← inner product between rows.

Result 2A.13. A matrix $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^{-1} = \mathbf{A}'$. For an orthogonal matrix, $\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}$, so the columns are also mutually perpendicular and have unit lengths.

$\hookrightarrow$ inner product between columns.

- determinant of a square matrix

Definition 2A.24. The *determinant* of the square $k \times k$ matrix $\mathbf{A} = \{a_{ij}\}$, denoted by $|\mathbf{A}|$, is the scalar

$$|\mathbf{A}| = a_{11} \qquad \text{if } k = 1$$

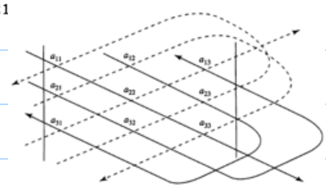
$$|\mathbf{A}| = \sum_{j=1}^{k} a_{1j}|\mathbf{A}_{1j}|(-1)^{1+j} \qquad \text{if } k > 1$$

where $\mathbf{A}_{1j}$ is the $(k-1) \times (k-1)$ matrix obtained by deleting the first row and jth column of $\mathbf{A}$. Also, $|\mathbf{A}| = \sum_{j=1}^{k} a_{ij}|\mathbf{A}_{ij}|(-1)^{i+j}$, with the ith row in place of the first row.

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}(-1)^2 + a_{12}a_{21}(-1)^3 = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}(-1)^2 + a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}(-1)^3 + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}(-1)^4$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{32}a_{23}a_{11}$$



Result 2A.11. Let $\mathbf{A}$ and $\mathbf{B}$ be $k \times k$ square matrices.

(a) $|\mathbf{A}| = |\mathbf{A}'|$

(b) If each element of a row (column) of $\mathbf{A}$ is zero, then $|\mathbf{A}| = 0$

(c) If any two rows (columns) of $\mathbf{A}$ are identical, then $|\mathbf{A}| = 0$

geometric interpretation of determinant ↔ eigensystem

(d) If $\mathbf{A}$ is nonsingular, then $|\mathbf{A}| = 1/|\mathbf{A}^{-1}|$; that is, $|\mathbf{A}||\mathbf{A}^{-1}| = 1$.

(e) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ $1 = |I| = |AA^{-1}| = |A||A^{-1}|$

(f) $|c\mathbf{A}| = c^k|\mathbf{A}|$, where c is a scalar.

- eigenvalues and eigenvectors of a square matrix

Definition 2A.30. Let $\mathbf{A}$ be a $k \times k$ square matrix and $\mathbf{I}$ be the $k \times k$ identity matrix. Then the scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ satisfying the polynomial equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ are called the *eigenvalues* (or *characteristic roots*) of a matrix $\mathbf{A}$. The equation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ (as a function of λ) is called the *characteristic equation*.

$$\mathbf{A} = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}$$

three roots: $\lambda_1 = 9$, $\lambda_2 = 9$, and $\lambda_3 = 18$

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 13-\lambda & -4 & 2 \\ -4 & 13-\lambda & -2 \\ 2 & -2 & 10-\lambda \end{vmatrix} = -\lambda^3 + 36\lambda^2 - 405\lambda + 1458 = 0$$

- For general $\mathbf{A}$, eigenvalues could be real or complex values
  - every eigenvalue of symmetric matrix is real