

In the following, vectors are always *column* vectors.

- 1. Basis:** $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for a vector space V if its vectors are linearly independent and span V , i.e. that for all \mathbf{v} in V there exists a unique decomposition $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{b}_i$. If V possesses a finite basis B , then the number of vectors in B is called the **dimension** of B .
- 2. Row space of an $m \times n$ matrix A :** $\{\mathbf{b}^t | \mathbf{b}^t = \mathbf{x}^t A \text{ for some } \mathbf{x} \text{ in } R^m\}$, i.e. \mathbf{b}^t is a linear combination of the rows of A .
- 3. Column space of an $m \times n$ matrix A :** $\{\mathbf{b} | \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \text{ in } R^n\}$. This is the same as the row space of A^t .

The dimensions of the row and column space are always equal, and this common dimension is called the **rank** of the matrix A .

4. Inner product, norm, distance

inner product

- $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
- $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
- $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$
- $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$
- $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$

norm

- $\|\mathbf{u}\| \geq 0$
- $\|\mathbf{u}\| = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$
- $\|k\mathbf{u}\| = |k| \|\mathbf{u}\|$
- $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

distance

- $d(\mathbf{u}, \mathbf{v}) \geq 0$
- $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
- $d(\mathbf{u}, \mathbf{v}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{v}$
- $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$

To each inner product, one can associate a norm $\|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{v} \rangle^{1/2}$. To each norm one can associate a distance $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.

If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ we say that \mathbf{u} is **orthogonal** to \mathbf{v} and we write $\mathbf{u} \perp \mathbf{v}$. Every set of orthogonal (nonzero) vectors is independent.

- 5. Cauchy-Schwarz inequality:** $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$.
- 6. Orthonormal basis:** $B = \{\mathbf{b}_1 | \dots | \mathbf{b}_n\}$ is orthonormal if

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \begin{cases} 1 & \text{if } i = j \text{ (}\|\mathbf{b}_i\| = 1\text{);} \\ 0 & \text{otherwise (}\mathbf{b}_i \perp \mathbf{b}_j \text{ if } i \neq j\text{).} \end{cases}$$

The (column) basis vectors of an orthonormal basis of R^n form an **orthogonal** matrix $Q = (\mathbf{b}_1 | \dots | \mathbf{b}_n)$. The transpose Q^t of an orthogonal matrix is orthogonal, and we have: $Q^t Q = I$. It is very easy to express a vector \mathbf{v} with respect to an orthonormal basis B . Observe that $Q^{-1} = Q^t$! We have :

$$\mathbf{v} = \sum_1^n \langle \mathbf{v}, \mathbf{b}_i \rangle \mathbf{b}_i,$$

i.e. that $[\mathbf{v}]_B = [\langle \mathbf{v}, \mathbf{b}_1 \rangle, \dots, \langle \mathbf{v}, \mathbf{b}_n \rangle]^t_B$. It is also easy to find the projection of a vector $\mathbf{v} \in V$ onto a subspace W having an orthonormal basis $B' = \{\mathbf{b}'_1, \dots, \mathbf{b}'_m\}$:

$$\text{Proj}_W \mathbf{v} = \sum_{i=1}^k \langle \mathbf{v}, \mathbf{b}'_i \rangle \mathbf{b}'_i.$$

If B is an orthonormal basis for V , $[\mathbf{v}]_B = [\alpha_1, \dots, \alpha_n]^t$ and $[\mathbf{w}]_B = [\beta_1, \dots, \beta_n]^t$, then we have:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n \alpha_i \beta_i, \quad \|\mathbf{v}\|^2 = \sum_{i=1}^n \alpha_i^2, \quad d(\mathbf{v}, \mathbf{w}) = \left(\sum_{i=1}^n (\alpha_i - \beta_i)^2 \right)^{1/2}.$$

Remark. Given a vector space V , a subspace W of V and a vector $\mathbf{v} \in V$, \mathbf{v} can be decomposed as $\mathbf{v} = \mathbf{p} + \mathbf{w}$ where \mathbf{p} is the projection of \mathbf{v} onto W and \mathbf{w} a vector orthogonal to W . This decomposition is unique. Furthermore the projection \mathbf{p} is the **best approximation** of \mathbf{v} by a vector in W .

15. Eigenvalues and eigenvectors: λ is an **eigenvalue** and \mathbf{v} an **eigenvector** for the transformation T from V to V if $T(\mathbf{v}) = \lambda\mathbf{v}$. If V has finite dimension, T is uniquely determined by its matrix representation in some basis B , and the set of eigenvalues is independent of the choice of the basis B . Therefore we will restrict ourselves to the study of linear transformations represented by square matrices. Eigenvalues and eigenvectors of A must satisfy the equation: $A\mathbf{x} = \lambda\mathbf{x}$. A number λ is an eigenvalue for a square matrix A of order n if and only if it is a root of the **characteristic polynomial** (of degree n): $\det(A - \lambda I) = 0$. The n roots of the characteristic polynomial need not be distinct or even real.

Corresponding to each eigenvalue λ is the **eigenspace** of corresponding eigenvectors.

Since $A^n\mathbf{x} = \lambda^n\mathbf{x}$, the eigenvalues of A^n are the n th powers of the eigenvalues of A . In particular (show it!) the eigenvalues of A^{-1} are the inverses of the eigenvalues of A .

16. Diagonalization: A matrix A is **diagonalizable** if there exists a matrix P such that $P^{-1}AP$ is diagonal, i.e. A is similar to a diagonal matrix D . A is diagonalizable (theorem) if and only if it has n independent eigenvectors. The columns of P correspond to the n eigenvectors. If A has n distinct eigenvalues it is diagonalizable, since the eigenvectors associated with distinct eigenvalues are independent. However, this condition is not necessary (consider the identity matrix I).

Let $T(\mathbf{x}) = A\mathbf{x}$ be a linear transformation from R^n to R^n and A its matrix representation in the standard basis of R^n . Assume A has n independent eigenvectors $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and consider the basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. Let $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{b}_i$, i.e. $[\mathbf{v}]_B = [\alpha_1 \ \dots \ \alpha_n]_B^t$. We have:

$$T(\mathbf{v}) = \sum_{i=1}^n \alpha_i T(\mathbf{b}_i) = \sum_{i=1}^n \alpha_i \lambda_i \mathbf{b}_i$$

or $[T(\mathbf{v})]_B = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}_B [\mathbf{v}]_B$.

In the basis B , the transformation has a diagonal matrix representation!

17. Orthogonal diagonalization: If there exists an orthogonal matrix P such that $P^tAP = D$ (diagonal matrix), then we say that A is orthogonally diagonalizable. This occurs when A has an orthonormal set of n eigenvectors. A necessary and sufficient condition is that A be symmetric (theorem), i.e. $A^t = A$. In particular the eigenvalues of a symmetric matrix are all *real*.

If A is symmetric, the eigenvectors associated with distinct eigenvalues are orthogonal (theorem). To find an orthonormal basis of eigenvectors it is therefore sufficient to Gram-Schmidtize each eigenspace!

18. Differential equations: Consider the system of differential equations

$$y'_i = \sum_{j=1}^n a_{ij} y_j \quad i = 1, \dots, n$$

with initial conditions: $y_i(0) = c_i \ i = 1, \dots, n$ or, in matrix form:

$$Y' = AY \quad Y(0) = C \quad \text{with the obvious correspondances.}$$

Assume A is symmetric, hence diagonalizable, and write: $P^{-1}AP = D$ or $A = PDP^{-1}$. The system becomes: $Y' = (PDP^{-1})Y$ or $P^{-1}Y' = DP^{-1}Y$. Let us perform the change of variable: $U = P^{-1}Y$. Then: $U' = P^{-1}Y'$ and the system simplifies to: $U' = DU$ where D is a diagonal matrix whose entries are the eigenvalues $\lambda_1, \dots, \lambda_n$ of A . The system is now separable, i.e. $u'_i = \lambda u_i \ i = 1, \dots, n$ and its general solution is given by:

$$u_i(x) = d_i e^{\lambda_i x} \quad i = 1, \dots, n.$$

Next: $Y(x) = PU(x)$ is the general solution for Y . The constants d_i are then determined by solving the linear system $Y(0) = C$. Alternatively one could have derived the particular solution in Y from the particular solution in U .

19. Fourier series: We want to approximate continuous functions on the interval $[0, 2\pi]$ by a **trigonometric polynomial**

$$t(x) = \frac{a_0}{2} + \sum_{k=1}^n [c_k \cos kx + d_k \sin kx].$$

Let V denote the vector space of continuous functions on $[0, 2\pi]$ and W the subspace of trigonometric polynomials of order n , with inner product given by:

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x) dx.$$

The best approximation of $f \in V$ by a function $t \in W$ is the projection $\text{Proj}_W(f)$ and it minimizes the **mean square error** $\|f - g\|^2 = \int_0^{2\pi} [f(x) - g(x)]^2 dx$ for all functions g in W .

The set $B = \{1, \cos x, \dots, \cos nx, \sin x, \dots, \sin nx\}$ is a basis for W . Furthermore, all its members are orthogonal. Consequently it is extremely easy to orthonormalize. Let us Gram-Schmidtize it:

$$\|1\|^2 = \int_0^{2\pi} 1^2 dx = 2\pi \quad \|\cos kx\|^2 = \int_0^{2\pi} (\cos kx)^2 dx = \pi \quad \|\sin kx\|^2 = \int_0^{2\pi} (\sin kx)^2 dx = \pi$$

$$B' = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos x, \dots, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin x, \dots, \frac{1}{\sqrt{\pi}} \sin nx \right\}$$

$$\begin{aligned} \text{Proj}_W f &= \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle \frac{1}{\sqrt{2\pi}} + \sum_{k=1}^n \left[\left\langle f, \frac{1}{\sqrt{\pi}} \cos kx \right\rangle \frac{1}{\sqrt{\pi}} \cos kx + \left\langle f, \frac{1}{\sqrt{\pi}} \sin kx \right\rangle \frac{1}{\sqrt{\pi}} \sin kx \right] \\ &= \frac{1}{2\pi} \langle f, 1 \rangle + \frac{1}{\pi} \sum_{k=1}^n [\langle f, \cos kx \rangle \cos kx + \langle f, \sin kx \rangle \sin kx] \\ &= \frac{1}{2} a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \end{aligned}$$

The numbers a_0 , a_k and b_k are the **Fourier coefficients** of f and they are given by the formulae:

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx \quad a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx.$$

As n grows, the approximation gets better; indeed, under weak conditions on f :

$$f(x) = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

THE END...