MATH 307

Ruiheng Su 2020

Linear systems

$$\underline{A}\underline{x} = \underline{b}$$

For $\underline{A} \in \mathbb{R}^{n\times m}$, we can write the i^{th} column of $\underline{A}$ as $\underline{A}$ times the i^{th} basis vector in R^m. $\underline{A}$ is a linear transformation from $\mathbb{R}^m \to \mathbb{R}^n$.

Row-echelon form

1. interchange two rows (pivoting)
2. scale one row by a non-zero number
3. add multiple of one row to another

Every matrix can be transformed to row-echelon form via elementary row operations.

Rank

Rank of matrix $\underline{A}$ Is the number of nonzero rows in the row-echelon form of $\underline{A}$.

- The column rank of $\underline{A}$ is $\dim(R(\underline{A}))$
- The row rank of $\underline{A}$ is $\dim(R(\underline{A}^\intercal))$

$$\text{rank}(\underline{A}) = \text{rank}(\underline{A}^\intercal)$$

Implies that the column rank is equal to the row rank

Solutions to a linear system

Let $[\underline{A}\underline{b}]$ be the augmented matrix of the system $\underline{A}\underline{x} = \underline{b}$. The system has

$$\begin{cases} \text{no solution} & \text{if rank}(\underline{A}) < \text{rank}([\underline{A}\underline{b}]) \\ \text{unique solution} & \text{if rank}(\underline{A}) = \text{rank}([\underline{A}\underline{b}]) \\ \infty & \text{if rank}(\underline{A}) > \text{rank}([\underline{A}\underline{b}]) \end{cases}$$

Consider the determinant of the square matrix $\underline{A}$. $\underline{A}\underline{x} = \underline{b}$ has the following number of solutions

$$\begin{cases} 0/\infty & \text{if } \det(\underline{A}) = 0 \\ \text{unique solution} & \text{if} \det(\underline{A}) \neq 0 \end{cases}$$

The inverse of $\underline{A}$ only exists if $\det(\underline{A}) \neq 0$.

The determinant of any diagonal square matrix is the product of its diagonal entires.

Consider a 2×2 matrix.

$$\underline{F} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The inverse of $\underline{F}$ is given by the formula:

$$\underline{F}^{-1} = \frac{1}{\det(\underline{F})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Linear independence

$\underline{v}_1, \ldots, \underline{v}_n$ is linearly independent if the only combination that gives the zero vector is

$$a\underline{v}_1 + b\underline{v}_2 + \cdots = \underline{0}$$

for all $a, b, \cdots = 0$.

The columns of a matrix $\underline{A}$ are linearly independent when the only solution to $\underline{A}\underline{x} = \underline{0}$ is $\underline{x} = \underline{0}$

LU and Cholesky decompositions

Any matrix $\underline{E}$ obtained from the identity matrix $\underline{I}$ by elementary row operations is an elementary matrix.

$\underline{E}\underline{A}$ applies the elementary row operations applied to $\underline{I}$ to get to $\underline{E}$ to $\underline{A}$.

Unit and atomic lower triangular matrix

Let $*$ is any number, the unit lower triangular matrix is a $m \times m$ matrix:

$$\underline{L} = \begin{bmatrix} 1 & & \\ * & 1 & \\ * & * & 1 \end{bmatrix}$$

Atomic lower triangular matrix, $\underline{L}_k$ for a given k, is an identity $\underline{I}$ matrix except entries L_{ik} for $i = k+1, \ldots, m$ are nonzero.

Let $\underline{A}$ be a $m \times n$ matrix. For a some given k, $\underline{L}_k\underline{A}$ is obtained by adding $L_{ik}A_{kj}$ to A_{ij} for each $i = k+1, \ldots, m$, $j = 1, \ldots, n$.

The idea is that each $\underline{L}_k$ is a matrix that "annihilates" non zero entries below the pivot. $\underline{L}_k$ is an elementary matrix.

The inverse $\underline{L}_k^{-1}$ (an atomic lower triangular matrix) is the same as $\underline{L}_k$ for all entries except the nonzero entries are multiplied with -1.

Multiplying $m\times m$ $\underline{L}_k$ for $k = 1, 2, \ldots, m-1$ results in a $m\times m$ unit triangular matrix with nonzero entries below the diagonal at the column k equal to the the nonzero entries below the diagonal for $\underline{L}_k$.

If $\underline{A}$ can be reduce by Gaussian elimination to row-echelon form without pivoting, then the LU decomposition of $\underline{A}$ is

$$\underline{A} = \underline{L}\underline{U},$$

where $\underline{L}$ is an unit lower triangular matrix and $\underline{U}$ is an upper triangular matrix.

LU decomposition algorithm

Einstein summation is not implied. For $m \times n$ matrix $\underline{A}_k$, we denote its entries as A^k_{ij}, where $k = 1, \ldots n-1$. Construct $\underline{L}_k$ so that L_{ik} for $i = k+1, \ldots, m$

$$L_{ik} = \frac{-A^k_{ik}}{A^k_{kk}}$$

The algorithm:

```
L = 1
U = 0
A_k = Matrix(...)
for k in range(0, n-1):
    """
    for some k, compute entires of L_k
    depending on A_k
    """
    L_k = ...
    A_k = L_k*A_k
    L = L*L_k.inverse()
U = A_k
```

Applications

Given $\underline{A} = \underline{L}\underline{U}$, we can solve $\underline{A}\underline{x} = \underline{b} \implies \underline{L}\underline{U}\underline{x} = \underline{b}$ by solving

1. $\underline{L}\underline{y} = \underline{b}$ for $\underline{y}$
2. $\underline{U}\underline{x} = \underline{y}$ for $\underline{x}$

The rank and determinant of $\underline{A}$ and $\underline{U}$ are equal. The inverse of $m \times n$ matrix $\underline{A}$ is

$$\underline{A}^{-1} = \begin{bmatrix} \underline{x}^1 \cdots \underline{x}^n \end{bmatrix}$$

for $\underline{x}^i$, $i = 1, \ldots, n$ obtained from solving $\underline{A}\underline{x}^i = I_{ij}$ for $i = 1, \ldots, n$, where $\underline{I}$ is the identity.

For $m \times n$ matrix $\underline{A}$, if $\text{rank}(\underline{A}) = m$ and $\underline{A}$ has an LU decomposition, then the matrices $\underline{L}$ and $\underline{U}$ are unique.

Positive definite and symmetric

Square matrix $\underline{A}$ is symmetric if $\underline{A}^\intercal = \underline{A}$. $\underline{A}$ is positive definite if

$$\underline{x}^\intercal \underline{A}\underline{x} > 0$$

for all column vector $\underline{x} \neq \underline{0}$. Equivalently, a symmetric matrix is positive definite if all its eigenvalues are positive.

Cholesky decomposition: for symmetric and positive definite square matrix $\underline{A}$, $\exists$ a lower triangular matrix $\underline{L}$ with positive diagonal entire (does not have to be 1) such that

$$\underline{A} = \underline{L}\underline{L}^\intercal$$

For a Cholesky decomposition of $\underline{A}$, $\exists$ a unit lower triangular matrix $\underline{L}$ and diagonal matrix $\underline{D}$ with positive entries such that

$$\underline{A} = \underline{L}\underline{D}\underline{L}^\intercal$$

Let $\sqrt{\underline{D}}$ be $\underline{D}$ when the squareroot is applied to its entries. Then

$$A = \left(\underline{L}\sqrt{\underline{D}}\right)\left(\sqrt{\underline{D}}\underline{L}^\intercal\right)$$

is also the Cholesky decomposition of A.

Cholesky decomposition algorithm

1. Reduce $\underline{A}$ to row-echelon form without interchanging or scaling.
2. Factor out the diagonal entries of $\underline{A}$ into a matrix $\underline{D}$. To do this, we need to find $\underline{J}$, where $D_{ij}J_{jk} = A_{ij}$.
3. $\underline{L} = \underline{J}^\intercal\sqrt{\underline{D}}$
4. $\underline{A} = \underline{L}\underline{L}^\intercal$

LU decomposition with partial pivoting

Let $\underline{L}_k$ be a $m \times m$ atomic lower triangular matrix. $\underline{P}$ be a $m \times m$ permutation matrix that switches rows i with j for $i, j > 0$. $\underline{P}\underline{L}_k\underline{P}$ is identical to $\underline{L}_k$ except L_{ik} and L_{jk} are switched.

Permutation matrix

Any $m \times m$ matrix $\underline{P}$ to swap two columns is the identity with those two column swapped. $\underline{P}$ are elementary matrices.

$$\underline{P}^\intercal = \underline{P}^{-1} = \underline{P}$$

$\underline{P}$ is symmetric, since $\underline{P}\underline{P} = I$, $\underline{P}\underline{P}^{-1} = I$, $\underline{P}\underline{P}^\intercal = I$.

LU decomposition with partial pivoting: any $\underline{A}$, $\exists$ a permutation matrix $\underline{P}$, unit lower triangular matrix $\underline{L}$ and upper triangular matrix $\underline{U}$ such that

$$\underline{A} = \underline{P}\underline{L}\underline{U}$$

The algorithm:

1. For each column, if the absolute value of the pivot is less than the absolute value of some entry below the pivot, find $\underline{P}$ to switch the row containing the bigger entry to the row of pivot
2. Move to the next column if all entries below the pivot are zero
3. Find annihilation matrix $\underline{L}$ for that column. Compute $\underline{L}\underline{A}$

In the end, we have

$$\underline{L}_n\underline{P}_n \cdots \underline{L}_1\underline{P}_1\underline{A} = \underline{U}$$

We define $\underline{P}$:

$$\underline{P} = \underline{P}_1 \cdots \underline{P}_n$$

and $\underline{L}$:

$$\underline{L} = \left(\underline{P}_1 \cdots \underline{P}_n\right)^{-1}\left(\underline{L}_n\underline{P}_n \cdots \underline{L}_1\underline{P}_1\right)^{-1}$$

Matrix norms and the condition number

A $\mathbb{R}^n$ norm is a function $\|\cdot\|$ such that

1. $\|\underline{x}\| \geq 0 \forall x \in \mathbb{R}^n$
2. $\|\underline{x}\| = 0$ iff $\underline{x} = \underline{0}$
3. $\|c\underline{x}\| = |c|\|\underline{x}\| \geq 0 \forall c \in \mathbb{R}, x \in \mathbb{R}^n$
4. $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\| \forall x, \underline{y} \in \mathbb{R}^n$ (triangular inequality)

The p-norm:

$$\|x\|_p = \left[\sum_{k=1}^{n} |x_k|^p\right]^{1/p}$$

As $p \to \infty$, the infinity norm for $\underline{x} \in \mathbb{R}^n$ is

$$\|\underline{x}\|_\infty = \max\left\{|x_k|\right\}, k = 1, 2, \ldots, n$$

In general, for any $\underline{x} \in \mathbb{R}^n$,

$$\|\underline{x}\|_1 \geq \|\underline{x}\|_2 \geq \|\underline{x}\|_\infty$$

Each differs by some constant.

Matrix norm

For some vector norm, the matrix norm measures the maximum stretching the matrix does to any vector. (For some vector $\underline{x} \in \mathbb{R}^n$, multiplication by a $m \times n$ matrix $\underline{x}$ is a linear transformation on $\underline{x}$.)

$$\|\underline{A}\| = \max_{\underline{x}\neq\underline{0}} \frac{\|\underline{A}\underline{x}\|}{\|\underline{x}\|} = \max_{\underline{x}\neq\underline{0}} \left\|\underline{A}\frac{\underline{x}}{\|\underline{x}\|}\right\| = \max_{\|\underline{x}\|=1} \|\underline{A}\underline{x}\|$$

A matrix norm corresponding to a vector norm satisfies

1. $\|\underline{A}\| \geq 0 \forall \underline{A} \neq \underline{0}$
2. $\|\underline{A}\| = 0$ iff $\underline{A} = \underline{0}$
3. $\|c\underline{A}\| = |c|\|A\| \forall c \in \mathbb{R}$
4. $\|\underline{A} + \underline{B}\| \leq \|\underline{A}\| + \|\underline{B}\|$
5. $\|\underline{A}\underline{B}\| \leq \|\underline{A}\|\|\underline{B}\|$
6. $\|\underline{A}\underline{x}\| \leq \|\underline{A}\|\|\underline{x}\| \forall \underline{x} \in \mathbb{R}^n$

For a $m \times m$ diagonal matrix $\underline{D}$, $\|\underline{D}\| = \max_i |D_{ii}|$ for $i = 1, \ldots, m$.

So even if $\|\underline{A}\|_2 = 1$, does not mean that $\underline{A} = \underline{I}$.

Condition number

The condition number of a non-singular (invertible) matrix $\underline{A}$ is

$$\text{cond}(A) = \|\underline{A}\|\|\underline{A^{-1}}\|$$

$\text{cond}(A) = \infty$ if $\underline{A}$ is non-invertible. The condition number is the ratio of the maximum and minimum stretching factors of $\underline{A}$

For the linear system $\underline{A}\underline{x} = \underline{b}$, if a small change $\Delta\underline{b}$ corresponds to a change $\Delta\underline{x}$ such that $\underline{A}(\underline{x} + \Delta\underline{x}) = \underline{b} + \Delta\underline{b}$, then

$$\frac{\|\Delta\underline{x}\|}{\|\underline{x}\|} \leq \text{cond}(A)\frac{\|\Delta\underline{b}\|}{\|\underline{b}\|}$$

For $\underline{b} \in \mathbb{R}^n$, the relative change (relative error) in $\underline{b}$ due to small change $\Delta\underline{b}$ is

$$\frac{\|\Delta\underline{b}\|}{\|b\|}$$

So the condition number multiplied by the relative change in $\underline{b}$ gives an upper bound on the relative error in the solution $\underline{x}$.

Polynomial interpolation

Given $d + 1$ points $(t_0, y_0), \ldots (t_d, y_d)$, polynomial interpolation using the monomial basis wants to find a polynomial of degree at most equal to d, of the form

$$p(t) = c_0 + c_1 t + \cdots + c_d t^d$$

such that $p(t_k) = y_k$ for each $k = 0, \ldots d$.

Imposing an additional condition such as restricting $p(t_1) = 0$, increases the unique polynomial order by 1.

Monomial basis

The elements $1, t, \ldots, t^d$ form the monomial basis of the vector space, $\mathbb{P}_d$, of all polynomials of degree less than or equation to d. To find the coefficients, we solve the following linear systems of equations for the coefficients $c_0, \ldots c_d$.

$$\begin{aligned} c_0 + c_1 t_0 + \cdots + c_d t_0^d &= y_0 \\ c_0 + c_1 t_1 + \cdots + c_d t_1^d &= y_1 \\ &\vdots \\ c_0 + c_1 t_d + \cdots + c_d t_d^d &= y_d \end{aligned}$$

In matrix vector form:

$$\underline{Ac} = \underline{y} = \begin{bmatrix} 1 & t_0 & \ldots & t_0^d \\ 1 & t_1 & \ldots & t_1^d \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_d & \ldots & t_d^d \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_d \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_d \end{bmatrix}$$

$\underline{A}$ is the Vandermonde matrix. cond($\underline{A}$) gets very large as size of the matrix grows.

> For $d+1$ data points, where $t_i \neq t_j$ for $i \neq j$, there exists a unique polynomial $p(t)$ of degree at most d that perfectly interpolates the data points. The Vandermonde matrix is invertible when the values t_k are distinct.
>
> The implication is that the monomial, Lagrange, and Newton representations produce the same results.

Lagrange basis

Given $d+1$ data points $(t_0, y_0), \ldots (t_d, y_d)$, Lagrange interpolation seeks a polynomial of the form

$$p(t) = c_0 \ell_0(t) + c_1 \ell_1(t) + \cdots + c_d \ell_d(t)$$

where the Lagrange basis $\ell_0(t), \ldots \ell_d(t)$ are given by

$$\ell_k(t) = \prod_{j=0, j\neq k}^{d} \frac{t - t_j}{t_k - t_j}$$

The Lagrange polynomials have the Kronecker delta property:

$$\ell_k(t_j) = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases} \text{ for } k, j = 0, \ldots, d$$

So $p(t_j) = c_j = y_j$, and

$$p(t) = y_0 \ell_0(t) + y_1 \ell_1(t) + \cdots + y_d \ell_d(t)$$

Newton basis

Given $d+1$ data points $(t_0, y_0), \ldots (t_d, y_d)$, Newton interpolation seeks a polynomial of the form

$$p(t) = c_0 p_0(t) + c_1 p_1(t) + \cdots + c_d p_d(t)$$

where the Newton basis $p_0(t), \ldots, p_d(t)$ ar give by

$$\begin{aligned} p_0 &= 1 \\ p_1 &= (t - t_0) \\ p_2 &= p_1(t - t_1) \\ &\vdots \\ p_{d-1} &= p_\cdot p_2 \ldots p_d(t - t_{d-1}) \end{aligned}$$

Cubic spine interpolation

Given $N+1$ data points, $(t_0, y_0), \ldots, (t_N, y_N)$, where $t_i \neq t_j$ for $i \neq j$, a cubic spline seeks a piecewise function $p(t)$ consisting of N cubic polynomials, each cubic polynomial has the form

$$p_k(t) = a_k(t - t_{k-1})^3 + b_k(t - t_{k-1})^2 + c_k(t - t_{k-1}) + d_k$$

Since each cubic polynomial requires 4 parameters, we need $4N$ parameters to interpolate N data points. We require

1. $p_k(t_{k-1}) = y_{k-1}$ for $k = 1, \ldots, N$
2. $p_k(t_k) = y_k$ for $k = 1, \ldots, N$
3. $p'_k(t_k) = p'_{k+1}(t_k)$ for $k = 1, \ldots N-1$
4. $p''_k(t_k) = p''_{k+1}(t_k)$ for $k = 1, \ldots N-1$

yielding $4N-2$ conditions. The last condition is that the spline has zero second derivatives at the end points (natural cubic spline).

> The entries of the coefficient matrix
>
> $$C = \begin{bmatrix} a_1 & \ldots & a_N \\ b_1 & \ldots & b_N \\ c_1 & \ldots & c_N \\ d_1 & \ldots & d_N \end{bmatrix}$$
>
> where $d_k = y_{k-1}$ for $k = 1, \ldots, N$, and a_k, b_k, c_k can be found by solving:
>
> $$\left[\begin{array}{c|c|c|c|c} A(L_1) & B & & & \\ \hline & A(L_2) & B & & \\ \hline & & \ddots & \ddots & \\ \hline & & & A(L_{N-1}) & B \\ \hline T & & & & V \end{array}\right] \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ \vdots \\ a_N \\ b_N \\ c_N \end{bmatrix} = \begin{bmatrix} y_1 - y_0 \\ 0 \\ 0 \\ \vdots \\ y_N - y_{N-1} \\ 0 \\ 0 \end{bmatrix}$$
>
> where $L_k = t_k - t_{k-1}$, and
>
> $$A(L) = \begin{bmatrix} L^3 & L^2 & L \\ 3L^2 & 2L & 1 \\ 6L & 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$
>
> $$T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad V = \begin{bmatrix} L_N^3 & L_N^2 & L_N \\ 0 & 0 & 0 \\ 6L_N & 2 & 0 \end{bmatrix}$$

The condition number of the cubic spline matrix does not increase drastically with the number of data points.

Finite difference method

The order of a differential equation is the highest order derivative in the equation.

Taylor series

For $f \in C^\infty$, the taylor series expansion of f about a is

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n$$

Finite difference formulas

f evaluated at a small distance Δx away of some point x is

$$\begin{aligned} f(x + \Delta x) &= f(x) + f'(x)\Delta x + \frac{f''(x)}{2!}\Delta x^2 + \ldots \\ f(x - \Delta x) &= f(x) - f'(x)\Delta x + \frac{f''(x)}{2!}\Delta x^2 + \ldots \end{aligned}$$

Rearranging the two series gives the forward and backward difference formulas:

$$\begin{aligned} f'(x) &= \frac{f(x + \Delta x) - f(x)}{\Delta x} + O(\Delta x) \implies \frac{f_{n+1} - f_n}{\Delta x} \\ f'(x) &= \frac{f(x) - f(x - \Delta x)}{\Delta x} + O(\Delta x) \implies \frac{f_n - f_{n-1}}{\Delta x} \end{aligned}$$

Subtracting $f(x + \Delta x)$ and $f(x - \Delta x)$ gives the central difference formula:

$$f'(x) = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} - \frac{(\Delta x)^2}{3!} f^3(x) + \ldots$$

$$f'(x) = \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} + O((\Delta x)^2) \implies \frac{f_{n+1} - f_{n-1}}{2\Delta x}$$

The central difference formula for second derivatives:

$$\begin{aligned} f(x + \Delta x) + f(x - \Delta x) &= 2f(x) + \frac{2(\Delta x)^2}{2!} f''(x) + \ldots \\ &= 2f(x) + (\Delta x)^2 f''(x) + \frac{2(\Delta x)^4}{4!} f^4(x) + \ldots \end{aligned}$$

we see that that the odd power terms are eliminated. We can rearrange and solve for $f''(x)$:

$$f''(x) = \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{(\Delta x)^2} + O((\Delta x)^2) \implies \frac{f_{n+1} - 2f_n + f_{n-1}}{(\Delta x)^2}$$

1. Discretize the domain into $N+1$ steps. Choose some N such that for $n = 1, \ldots, N$

   $$x_n = x_0 + n(x_N - x_0)/(N+1)$$

2. Substitute the finite difference equations into the differential equation

3. Form N linear equations with N unknowns. Write into matrix vector form
4. Account for Dirichlet boundary conditions ($f_0 = f(x_0) = \alpha$, $f_{N+1} = f(x_{N+1}) = \beta$)
5. Account for Neumann boundary conditions by approximating y_0 or y_{N+1} using forward/backward difference
6. Solve the linear system of equations for each f_n for $n = 1, \dots, N$ (since f_0 and f_{N+1} are known from boundary conditions)

Least squares approximation

Orthogonality

Vectors $\underline{x}, \underline{y} \in \mathbb{R}^n$ are orthogonal their dot product is the zero vector.

$$\underline{x} \cdot \underline{y} = \underline{x}^\intercal \underline{y} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ y_n \end{bmatrix} = \|\underline{x}\| \|\underline{y}\| \cos(\theta),$$

θ is the acute angle between the two vectors. Other properties

- $\underline{x} \cdot \underline{y} = \underline{y} \cdot \underline{x}$
- $\left(a\underline{x} + b\underline{y}\right) \cdot \underline{z} = a\left(\underline{x} \cdot \underline{z}\right) + b\left(\underline{y} \cdot \underline{z}\right)$
- $\underline{x} \cdot \underline{x} = \|\underline{x}\|^2$
- $\underline{Ax} \cdot \underline{y} = \underline{x} \cdot \underline{A}^\intercal \underline{y}$

(Cauchy-Schwartz inequality) For $\underline{x}, \underline{y} \in \mathbb{R}^n$, then

$$|\underline{x} \cdot \underline{y}| \leq \|\underline{x}\| \|\underline{y}\|$$

(Triangle inequality) For $\underline{x}, \underline{y} \in \mathbb{R}^n$, then

$$\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$$

A set of vectors $\underline{x}_1, \dots, \underline{x}_m$ are orthonormal if they are all unit vectors, and any vector $\underline{x}_i \perp \underline{x}_j$ in the set for $i \neq j$.

(Pythagoras theorem) For a set of orthogonal vectors, then

$$\|\underline{x}_1 + \dots + \underline{x}_m\| = \|\underline{x}\| + \dots + \|\underline{x}_m\|$$

Let $S_1 \subset \mathbb{R}^n$ and $S_2 \subset \mathbb{R}^n$ be subspaces. Then S_1 and S_2 are orthogonal if $\forall\ \underline{x} \in S_1$ and $\underline{y} \in S_2$,

$$\underline{x}_1 \cdot \underline{x}_2 = 0 \implies S_1 \perp S_2$$

(Orthogonal subspaces) Two subspaces are orthogonal if their basis vectors are orthogonal.

Let $B = [b_1, \dots, b_k]$ and $A = [a_1, \dots, a_\ell]$ be sets basis vectors for $S_B, S_A \in \mathbb{R}^n$. $S_B \perp S_A$ if

$$\underbrace{B^\intercal A}_{k \times \ell} = \underline{0}$$

The orthogonal complement of $S \subset W$ is the subspace

$$S^\perp = \{\underline{x} \in W : \underline{x} \perp \underline{y}\ \forall\ \underline{y} \in S\}$$

$S^\perp$ is also a subspace of W.

Let $S \subset \mathbb{R}^n$, in general

$$\dim(S) + \dim(S^\perp)$$

If $S, V \subset \mathbb{R}^n$, $S \perp V$, and $\dim(S) = m$, then

$$V \in S^\perp \implies \dim(V) \leq n - \dim(S)$$

(Orthogonal decomposition) Let $U \subset W$, any vector $\underline{x} \in W$ can be written uniquely as

$$\underline{x} = \underline{x}_U + \underline{x}_{U^\perp}$$

for $\underline{x}_U \in U$, $\underline{x}_{U^\perp} \in U^\perp$

- $\left(U^\perp\right)^\perp = U$
- If $B, B_\perp$ are basis' for $U, U^\perp \subset W$, then $B \cup B_\perp$ is a basis for W
- $W \neq U \cap U^\perp$

Fundamental spaces of a matrix

The **null space** of $\underline{A}$ is the set of vectors that are mapped to zero under the transformation $\underline{A}$.

$$N(\underline{A}) = \{x \in \mathbb{R}^n | \underline{Ax} = \underline{0}\}$$

$\underline{A}$ is invertible if $N(\underline{A}) = \{\underline{0}\}$.

If we know that 0 is an eigenvalue of $\underline{B}$, then $\underline{Bv} = 0\underline{v} = \underline{0}$ for the corresponding eigenvector $\underline{v} \neq 0$. So $v \in N(\underline{B})$, and hence $\underline{B}$ is not invertible.

- $\dim(N(\underline{A}))$ is the nullity of A
- $\dim(N(\underline{A})) = n - \text{rank}(\underline{A})$

For any matrix $\underline{A}$,

$$N(\underline{A}) = N(\underline{A}^\intercal \underline{A})$$

To find the ,nullspace of $\underline{A}$

1. Row reduce $\underline{A}$ to find the pivots of $\underline{A}$
2. The non-pivots are the columns of the free parameters

The **column space** (also called range) of $\underline{A}$ is span of the column vectors of $\underline{A}$, equivalently,

$$R(\underline{A}) = \{\underline{Ax} | x \in \mathbb{R}^n\}$$

A basis for the column space are the pivot columns of $\underline{A}$.

We need to know $\underline{A}$ to find the basis.

- $\dim(R(\underline{A}))$ is the rank of $\underline{A}$

The **left null space** of $\underline{A}$ is

$$N(\underline{A}^\intercal)\{y \in \mathbb{R}^m | \underline{A}^\intercal \underline{y} = \underline{0}\}$$

The **row space** pf $\underline{A}$ is the set of vectors spanned by the rows of $\underline{A}$, equivalently,

$$R(\underline{A}^\intercal) = \{\underline{A}^\intercal \underline{y} | y \in \mathbb{R}^m\}$$

The row space of $\underline{A}$ is spanned by the non-zero rows in $\text{rref}(\underline{A})$.

(Rank-nullity) For $\underline{A} \in \mathbb{R}^{n \times m}$,

$$\dim(R(\underline{A})) + \dim(N(\underline{A})) = m$$

Similarly, for $\underline{A}^\intercal \in \mathbb{R}^{m \times n}$,

$$\dim(R(\underline{A}^\intercal)) + \dim(N(\underline{A}^\intercal)) = n,$$

n is also the number of rows of $\underline{A}$.

Let $\underline{A} \in \mathbb{R}^{m \times n}$, then

- $N(\underline{A}) = \left[R(\underline{A}^\intercal)\right]^\perp$
- $N(\underline{A}^\intercal) = \left[R(\underline{A})\right]^\perp$

Asking whether there exists an $\underline{x}$ such that $\underline{Ax} = \underline{b}$, for all $\underline{b}$ is the same as asking whether $\underline{b} \in R(\underline{A})$.

Orthogonal Projection

The projection of a vector $\underline{x}$ onto a vector $\underline{u}$ is

$$\text{proj}_{\underline{u}}(\underline{x}) = \frac{\underline{x} \cdot \underline{u}}{\|\underline{u}\|} \frac{\underline{u}}{\|\underline{u}\|}$$

We can write projections as matrix multiplication:

$$\text{proj}_{\underline{u}}(\underline{x}) = \underline{Px} \qquad \text{where } \underline{P} = \frac{1}{\|\underline{u}^2\|} \underline{uu}^\intercal$$

$\underline{P}$ is an orthogonal projection matrix with rank 1.

Gram-Schmidth orthogonalization algorithm

Given a basis $\{\underline{u}_1, \dots, \underline{u}_m\}$ for $S \subset \mathbb{R}^n$, the GS orthogonalization algorithm constructs an orthogonal basis of U, where

$$\begin{aligned}
\underline{v}_1 &= \underline{u}_1 \\
\underline{v}_2 &= \underline{u}_2 - \text{proj}_{\underline{v}_1} \underline{u}_2 \\
\underline{v}_3 &= \underline{u}_3 - \text{proj}_{\underline{v}_1} \underline{u}_3 - \text{proj}_{\underline{v}_2} \underline{u}_3 \\
&\vdots \\
\underline{v}_m &= \underline{u}_m - \text{proj}_{\underline{v}_1} \underline{u}_m - \cdots - \text{proj}_{\underline{v}_{m-1}} \underline{u}_m
\end{aligned}$$

and $\{\underline{v}_1, \dots, \underline{v}_m\}$ is an orthogonal basis of $\underline{U}$.

Orthogonal projection to a subspace

Let $U \subset \mathbb{R}^n$ with an orthogonal basis $B = \{\underline{u}_1, \dots, \underline{u}_m\}$. The orthogonal projection of a vector $\underline{x}$ onto $\underline{U}$ is the sum of the projections of $\underline{x}$ onto each orthogonal basis vector of $\underline{U}$:

$$\text{proj}_U \underline{x} = \sum_{i=1}^{m} \frac{\underline{x} \cdot \underline{u}_i}{\|\underline{u}_1\|^2} \underline{u}_i$$

As a matrix multiplication:

$$\text{proj}_U \underline{x} = \underline{P}\underline{x} = \left[\sum_{i=1}^{m} \frac{\underline{u}_i \underline{u}_i^\intercal}{\|\underline{u}_1\|^2} \right] \underline{x}$$

$\underline{P}$ the **orthogonal projector** onto $\text{span}(B)$.

- $\underline{P}$ is a symmetric matrix
- $\underline{P} = \underline{P}^2$
- $\text{rank}(\underline{P}) = m$

The orthogonal projector $\underline{Q}$ onto $U^\perp$

$$\text{proj}_{U^\perp}(x) = \underline{Q}x,$$

is related to $\underline{P}$ in that:

$$\underline{Q} = \underline{I} - \underline{P}$$

(Projection theorem) Let $U \subset \mathbb{R}^n$, and let $\underline{x} \in \mathbb{R}^n$. Then,

$$\underline{x} - \text{proj}_U(\underline{x}) \in U^\perp$$

$\text{proj}_U(\underline{x})$ is the closest vector in U to $\underline{x}$:

$$\|\underline{x} - \text{proj}_U(\underline{x})\| < \|\underline{x} - \underline{y}\|$$

for all $\underline{y} \in U$, $\underline{y} \neq \text{proj}_U(\underline{x})$. ($\text{proj}_U(\underline{x})$ is the projection of $\underline{x}$ onto each orthogonal basis vector of U)

Let U be a m dimensional subspace of $\mathbb{R}^n$. The eigenvalues of $\underline{P}_U$ are 0, 1. $\lambda = 0$ has multiplicities equal to $n - m$. $\lambda = 1$ has multiplicities equal to m.

QR decomposition by Gram-Schmidt orthogonalization

The QR decomposition of a $n \times m$ matrix $\underline{A}$ with independent columns factors the matrix into an orthogonal matrix, $\underline{Q}$, and an invertible upper triangular matrix P.

By construction, $R(\underline{A}) = R(\underline{Q})$

Orthogonal matrix

A square matrix $\underline{A}$ is orthogonal if

$$\underline{A}\underline{A}^\intercal = \underline{A}^\intercal \underline{A} = I \implies \underline{A}^\intercal = \underline{A}^{-1}$$

If $\underline{A}$ is an orthogonal matrix, then

- $\|\underline{A}\underline{x}\| = \|\underline{x}\|$ for all $\underline{x} \in \mathbb{R}^n$.
- Columns of $\underline{A}$ are orthonormal
- Rows of $\underline{A}$ are orthonormal
- $\|\underline{A}\|_2 = 1$

QR decomposition

Let $\underline{A} \in \mathbb{R}^{n \times m}$,

1. Use Gram-Schmidt to construct an orthonormal basis of the column space
2. Put the orthonormal basis vectors into columns of a matrix $\underline{Q}$.
3. Solve for $\underline{R}_1$ by computing $\underline{Q}^\intercal \underline{A}$ (since $\underline{Q}$ is an orthogonal matrix)
4. Extend the basis into an orthonormal basis for $\mathbb{R}^n$, where the additional basis vectors are the orthonormal basis for $\left[R(\underline{A})\right]^\perp = N(\underline{A}^\intercal)$.
5. Add the new basis vectors as columns to $\underline{Q}$. Add rows of zeros to R_1 as needed
6. $\underline{A} = \underline{Q}\underline{R}$

The Gram-Schmidt algorithm is not the most efficient way to compute the QR decomposition.

QR decomposition by elementary reflectors

An elementary reflector (Householder transformation) is a matrix of the form

$$\underline{H} = \underline{I} - 2\frac{\underline{u}\underline{u}^\intercal}{\|\underline{u}\|^2}$$

for some non-zero vector $\underline{u} \in \mathbb{R}^n$.

Elementary reflectors are orthogonal matrices.

- The Householder transformation is also Hermitian

$\underline{H}$ reflects vectors through a hyperplane orthogonal to $\underline{u}$ (in other words, u is the normal vector to that plane).

- $\underline{H}\underline{v} = \underline{v}$ for all $v \in \text{span}(\underline{u})^\perp$ (the reflector maps all vectors on the hyperplane orthogonal to $\underline{u}$ onto themselves)

For some integer $k \leq n$ and a vector $\underline{a} \in \mathbb{R}^n$, partition the vector:

$$\underline{a} = \begin{bmatrix} \underline{a}_1 \\ \underline{a}_2 \end{bmatrix},$$

where $\underline{a}_1$ is $[a_1, \dots, a_{k-1}]^\intercal$, $\underline{a}_2 = [a_k, \dots, a_n]^\intercal$. Define the constant α as

$$\alpha = -\text{sign}(a_k)\|\underline{a}_2\|$$

and let

$$\underline{u} = \begin{bmatrix} \underline{0} \\ \underline{a}_2 \end{bmatrix} - \alpha \underline{e}_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_k - \alpha \\ a_{k+1} \\ \vdots \\ a_n \end{bmatrix}$$

and $\underline{H}$ be the corresponding elementary reflector of through $\underline{u}$,

$$\underline{H}\underline{a} = \begin{bmatrix} a_1 & \dots & a_{k+1} & \alpha & 0 & \dots & 0 \end{bmatrix}^\intercal$$

Let $\underline{A} \in \mathbb{R}^{n \times m}$, with $n > m$. There exists $\underline{H}_1, \dots, H_m$, such $\underline{H}_m \dots \underline{H}_1 \underline{A} = \underline{R}$, where $\underline{R}$ is upper triangular. Therefore

$$A = \underline{H}_1 \dots \underline{H}_m \underline{R} = \underline{Q}\underline{R}$$

$\underline{H}_k \underline{A}$ returns with the same size of $\underline{A}$ with entries on below the k row and k column annihilated.

Let $\underline{A} \in \mathbb{R}^{n \times m}$,

1. Using column k from $\underline{H}_{k-1} \dots \underline{H}_1 \underline{A}$, find $\underline{H}_k$
2. Then, $\underline{R} = \underline{H}_k \dots \underline{H}_1 \underline{A}$

Least squares approximation

Let $\underline{A}$ be an $m \times n$ matrix with $m > n$ and $\text{rank}(\underline{A}) = n$. The best approximation of the system

$$\underline{A}\underline{x} = \underline{b}$$

is the vector $\underline{z}$ that minimizes the distance $\|A\underline{z} - \underline{b}\|_2$. Since the two norm is used, this approximation is called the **least squares approximation**.

> Let $\underline{A}$ be an $m \times n$ matrix with $m > n$ and $\text{rank}(\underline{A}) = n$. The LSA to the system $\underline{A}\underline{x} = \underline{b}$ is the solution to the **normal equation**:
>
> $$\underline{A}^\intercal \underline{A}\underline{z} = \underline{A}^\intercal \underline{b}$$
>
> and $\underline{z}$ is unique.

$\underline{A}\underline{x} \in R(\underline{A})$. By the projection theorem, the closest vector in $R(\underline{A})$ to $\underline{b}$ is $\text{proj}_{R(\underline{A})}\underline{b}$. If

$$\underline{A}\underline{z} = \text{proj}_{R(\underline{A})}\underline{b}$$

Then $\text{proj}_{R(\underline{A})}\underline{b} - \underline{b} = \underline{A}\underline{z} - \underline{b} \in R(\underline{A})^\perp$. So

$$\underline{A}^\intercal(\underline{A}\underline{z} - \underline{b}) = 0 \implies \underline{A}^\intercal \underline{A}\underline{z} = \underline{A}^\intercal \underline{b}$$

Since $\text{rank}(\underline{A}) = n$, $\underline{A}$ is invertible, and there exists a unique solution to the normal equation.

When $\underline{A}$ is not invertible:

1. implies that the columns of $\underline{A}$ are linearly dependent
2. We can row-reduce $\underline{A}$ to find is pivot columns and a basis of $R(\underline{A})$
3. Write the basis as columns in a new matrix $\widetilde{\underline{A}}$; $R(\underline{A}) = R(\widetilde{\underline{A}})$
4. We know $\widetilde{\underline{A}}$ will be invertible
5. Follow through with the least squares procedure

.

> The least squares approximation to the system $\underline{A}\underline{x} = \underline{b}$ is the solution of the system
>
> $$\underline{R}_1\underline{z} = \underline{c}_1$$
>
> where
>
> $$\underline{A} = \underline{Q}\underline{R} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$
>
> and
>
> $$\underline{Q}^{-1}\underline{b} = \underline{Q}^\intercal \underline{b} = \begin{bmatrix} Q_1^\intercal \\ Q_2^\intercal \end{bmatrix} \underline{b} = \begin{bmatrix} \underline{c}_1 \\ \underline{c}_2 \end{bmatrix}$$

Proof. We want to minimize $\|A\underline{z} - \underline{b}\|_2$, which is the same as minimizing $\|A\underline{z} - \underline{b}\|_2^2$. Recall $\|Q\| = 1$ since Q is an orthogonal matrix. Then,

$$\begin{aligned} \|\underline{A}\underline{z} - \underline{b}\|^2 &= \|\underline{Q}\underline{R}\underline{z} - \underline{b}\|^2 \\ &= \|\underline{R}\underline{z} - \underline{Q}^\intercal \underline{b}\|^2 \\ &= \left\| \begin{bmatrix} \underline{R}_1\underline{z} \\ 0 \end{bmatrix} - \begin{bmatrix} \underline{c}_1 \\ \underline{c}_2 \end{bmatrix} \right\|^2 \\ &= \|\underline{R}_1\underline{z} - \underline{c}_1\|^2 + \|\underline{c}_2\|^2 \end{aligned}$$

Since $R(\underline{Q}_2) \subset R(\underline{A})^\perp$, then $\underline{c}_2 \in R(\underline{A})^\perp$, and does not depend on $\underline{z}$. So The minimum value occurs when $\underline{R}_1\underline{z} = \underline{c}_1$. □

Fitting models to data

For the linear case: given m data points, $(t_1.y_1), \ldots, (t_m, y_m)$, where $m \geq 2$, $t_i \neq t_j$ for $i \neq j$, the line

$$y = c_1 + c_2 t$$

that best fits the data forms a set of linear equations:

$$\begin{aligned} y_1 &= c_1 + c_2 t_1 \\ &\vdots \\ y_m &= c_1 + c_2 t_m \end{aligned}$$

written in matrix vector form:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} c_1 & c_2 \end{bmatrix} \implies \underline{y} = \underline{A}\underline{c}$$

The LSA approximation finds the vector $\underline{c}$ that best approximates the system $\underline{y} = \underline{A}\underline{c}$.

General least squares data fitting problem

Given m data points, where $t_i \neq t_j$ for $i \neq j$, and the model function

$$f(t, \underline{c}) = f(t, c_1, \ldots, c_n)$$

We want to minimize the sum of squared errors (SSE):

$$\sum_{i=1}^{m} (y_i - f(t_i, \underline{c}))^2$$

If f is a function of the form:

$$f(t, \underline{c}) = c_1 f_1(t) + \cdots + c_n f_n(t)$$

Then the problem is linear (f_n maybe nonlinear).
Assuming there are more data points than there are parameters ($m > n$), and the terms $f_1, \ldots, f_n$ are linearly independent (necessary so that $\text{rank}(\underline{A}) = n$), the least square approximation to the system

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} f_1(t_1) & \ldots & f_n(t_1) \\ \vdots & & \vdots \\ f_1(t_m) & \ldots & f_n(t_m) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

minimizes the SSE.

Eigenvalue problems

Eigenvalues and eigenvectors

An eigenvalue of a square matrix $\underline{A}$ is a number λ such that

$$\underline{A}\underline{v} = \lambda \underline{v}$$

for a corresponding eigenvector $\underline{v}$. This implies that

$$(\underline{A} - \lambda \underline{I})\underline{v} = \underline{0}$$

and $(\underline{A} - \lambda \underline{I})$ is not invertible. This is equivalent to

$$\det(\underline{A} - \lambda \underline{I}) = 0$$

The characteristic polynomial of a $n \times n$ matrix $\underline{A}$ is

$$\det(\underline{A} - \lambda \underline{I}) = c_{\underline{A}}(\lambda)$$

is an n^{th} degree polynomial whose roots are the n eigenvalues of $\underline{A}$.

> If λ is a non-zero eigenvalue of $\underline{A}$, then λ is also an eigenvalue of $\underline{A}^\intercal$.

Since $\underline{A}^\intercal$ and $\underline{A}$ have the same characteristic equation.

> λ is an eigenvalue of invertible matrix $\underline{A}$ then $1/\lambda$ is an eigenvalue of $\underline{A}^{-1}$

> Two **similar matrices** have the same eigenvalue.
>
> A matrix $\underline{B}$ is similar to $\underline{A}$ if there exists an invertible matrix $\underline{P}$ such that $\underline{B} = \underline{P}\underline{A}\underline{P}^{-1}$.

> If $\underline{A}$ is a real matrix with an eigenvalue-eigenvector pair λ, $\underline{v}$, then
>
> $$\overline{\lambda}, \overline{v}$$
>
> is also an eigenvalue-eigenvector pair of $\underline{A}$

Algebraic multiplicity

The number of times an eigenvalue appears as the root for characteristic polynomial.

Geometric multiplicity

The number of vectors in the eigenbasis for a particular eigenvalue.

> Geometric multiplicity ≤ Algebraic multiplicity

For a given $n \times n$ matrix $\underline{A}$, we can have

1. n distinct eigenvalues, each eigenvalue has AM = GM
2. For some $k < n$, $n - k$ distinct eigenvalues, each eigenvalue has $AM = GM$

3. For some $k < n$, $n - k$ distinct eigenvalues, each eigenvalue has GM < AM

> Eigenvectors from distinct eigenvalues are linearly independent.

Diagonalizability

$\underline{A}$ is **diagonalizable** if there exists an invertible matrix $\underline{P}$ and a diagonal matrix $\underline{D}$ such that $\underline{A} = \underline{PDP}^{-1}$.

> $\underline{A}$ is **diagonalizable** if the columns of $\underline{P}$ are eigenvectors and the diagonal entries of $\underline{D}$ are the corresponding eigenvalues of $\underline{A}$, where $\underline{A} = \underline{PDP}^{-1}$.

The multiplicity of an eigenvalue is the number of times it appears as a root to the characteristic polynomial.

> If a square matrix $\underline{A}$ has distinct eigenvalues (all eigenvalues have multiplicity 1), then $\underline{A}$ is diagonalizable.

> A square matrix $\underline{A}$ is diagonalizable iff every eigenvalue λ of multiplicity m yields m basic eigenvectors.
>
> (Every eigenvalue of algebraic multiplicity of m also has m geometric multiplicity. So some matrices with repeated eigenvalues are also diagonalizable.)

Always diagonalizable:

- Matrices with a basis of eigenvectors
- Matrices with distinct eigenvalues
- Hermitian matrices
- Non-zero orthogonal projection matrices

Sometimes diagonalizable:

- Matrices with repeated eigenvalues

Applications of diagonalization

Suppose $\underline{A} = \underline{PDP}^{-1}$.

1. Recall two properties of the determinant:

$$\det(\underline{BC}) = \det(\underline{B})\det(\underline{C})$$
$$\det(\underline{B}^{-1}) = 1/\det(\underline{B})$$

   Then,

$$\det(\underline{A}) = \det(\underline{P})\det(\underline{D})\det(\underline{P}^{-1}) = \det(\underline{D})$$

   so the determinant of a diagonalizable matrix is the product of its eigenvalues

2. The trace of a square matrix is the sum of its entries on the main diagonal. We have the following property

$$\text{tr}(\underline{AB}) = \text{tr}(\underline{BA})$$

   Therefore,

$$\text{tr}(\underline{A}) = \text{tr}(\underline{PDP}^{-1}) = \text{tr}(\underline{PP}^{-1}\underline{D}) = \text{tr}(\underline{D})$$

   So the trace of a diagonalizable matrix is the sum of its eigenvalues

3. The k^{th} power of a diagonalizable matrix, for $k \leq 0$, can be computed by

$$\underline{A}^k = \underline{PD}^k\underline{P}^{-1}$$

   If all eigenvalues of $\underline{A}$ are non-zero, then

$$\underline{D}^k = \begin{bmatrix} \lambda_1^k & 0 & & 0 \\ 0 & \lambda_2^k & & 0 \\ & & \ddots & \\ 0 & 0 & & \lambda_n^k \end{bmatrix}$$

   $\forall\, k \in \mathbb{Z}$.

> In general, even for non-diagonalizable matrices,
>
> $$\det(\underline{A}) = \lambda \ldots \lambda_n$$
> $$\text{tr}(\underline{A}) = \lambda_1 + \cdots + \lambda_n$$

Spectral theorem

Hermitian matrices

A square matrix $\underline{A}$ is Hermitian if $\underline{A}$ is equal to its own conjugate transpose

$$A_{ij} = \overline{A}_{ji} \implies \underline{A} = \overline{\underline{A}}^\intercal \implies \underline{A} = \underline{A}^*$$

- All diagonal entries of a Hermitian matrix are real
- A real valued matrix is Hermitian if it is symmetric.
- A Hermitian matrix is positive definite iff every eigenvalue is positive
- Having a basis of eigenvectors do not mean that the matrix is Hermitian

> If $\underline{A}$ is a real and symmetric matrix, then the eigenvalues of $\underline{A}$ are real numbers. (This is also true if $\underline{A}$ is Hermitian.)

Since $\underline{A}$ is real and symmetric,

$$\underline{A} = \underline{A}^\intercal = \overline{\underline{A}}^\intercal$$

where the over-line denotes the complex conjugate. Consider an eigenvalue, eigenvector pair $(\lambda, \underline{v})$:

$$\underline{v} \cdot (\overline{\underline{Av}}) = \underline{v} \cdot (\overline{\lambda \underline{v}}) = \overline{\lambda}\underline{v} \cdot \overline{\underline{v}} = \overline{\lambda}\|\underline{v}\|^2,$$

and,

$$\begin{aligned} \underline{v} \cdot (\overline{\underline{Av}}) &= (\overline{\underline{Av}}) \cdot \underline{v} \\ &= (\overline{\underline{Av}})^\intercal \underline{v} \\ &= \overline{\underline{v}}^\intercal (\overline{\underline{A}}^\intercal \underline{v}) \\ &= \overline{\underline{v}} \cdot (\underline{A}\overline{\underline{v}}) \qquad \text{since } \underline{A} = \underline{A}^\intercal \\ &= \overline{\underline{v}} \cdot \lambda \underline{v} \\ &= \lambda \|\underline{v}\|^2 \end{aligned}$$

It follows that $\lambda = \overline{\lambda}$, so λ is a real number.

> Let $\underline{A}$ be a real, symmetric matrix. And let λ_1 and λ_2 be distinct eigenvalues of $\underline{A}$. The corresponding eigenvectors $\underline{v}_1$ and $\underline{v}_2$ are orthogonal.

Proof. Consider two eigenvalues-eigenvector pair of $\underline{A}$, $\lambda_1 \neq \lambda_2$, $\underline{v}_1$, $\underline{v}_2$.

$$\underline{Av}_1 = \lambda_1 \underline{v}_1$$

Take the transpose of both sides, and multiply by $\underline{v}_1$.

$$\begin{aligned} (\underline{Av}_1)^\intercal \underline{v}_2 &= (\lambda_1 \underline{v}_1)^\intercal \underline{v}_2 \\ \underline{v}_1^\intercal \underline{A}^\intercal \underline{v}_2 &= \lambda_1 \underline{v}_1^\intercal \underline{v}_2 \\ \underline{v}_1^\intercal \lambda_2 \underline{v}_2 &= \lambda_1 \underline{v}_1^\intercal \underline{v}_2 \\ \lambda_2 \underline{v}_1 \cdot \underline{v}_2 &= \lambda_1 \underline{v}_1 \cdot \underline{v}_2 \end{aligned}$$

So $\underline{v}_1 \cdot \underline{v}_2 = 0$ since $\lambda_1 \neq \lambda_2$. □

> (Spectral theorem) Let $\underline{A}$ be a real symmetric matrix (or Hermitian matrix). There exists an orthogonal ("unitary") matrix $\underline{P}$ and diagonal matrix $\underline{D}$ such that $\underline{A} = \underline{PD}\overline{\underline{P}}^\intercal$ and $\underline{A}$ is diagonalizable ("orthogonally diagonalizable").
>
> The columns of $\underline{P}$ are a set of orthonormal eigenvectors, and the diagonal entries of $\underline{D}$ are the eigenvalues.

When $\underline{A}$ is Hermitian, $A = \underline{PD}\overline{\underline{P}}^\intercal$. We can write this matrix multiplication as

$$\underline{v}_1 \lambda_1 \overline{\underline{v}}_1^\intercal + \cdots + \underline{v}_n \lambda_n \overline{\underline{v}}_n^\intercal = \sum_{j=1}^{n} \lambda_j \underline{v}_j \overline{\underline{v}}_j^\intercal$$

Since $\underline{v}_j$ are of unit-norm, $\underline{v}_j \overline{\underline{v}}_j^\intercal$ is also the projection matrix onto $\underline{v}_j$. Then:

$$\underline{Aw} = \sum_{j=1}^{n} \lambda_j \underline{P}_j \underline{w}$$

$$\underline{A}^k = \sum_{j=1}^{n} \lambda_j^k \underline{P}_j$$

$$f(\underline{A}) = \sum_{j=1}^{n} f(\lambda_j) \underline{P}_j$$

Since $\underline{A}$ can be orthogonally diagonalized, $\|\underline{A}\| = \max_i |\lambda_i|$

Singular value decomposition

We want to generalize the idea of diagonalization to non-square matrices.

Consider a $m \times n$ real matrix $\underline{A}$.

1. Eigenvalues of $\underline{AA}^\top$ and $\underline{A}^\top\underline{A}$ are real and non-negative ($\lambda \geq 0$)
2. If λ is a non-zero eigenvalue of $\underline{AA}^\top$, then λ is an eigenvalue of $\underline{A}^\top\underline{A}$.
3. The spectral theorem says that both $\underline{AA}^\top$ and $\underline{A}^\top\underline{A}$ are orthogonally diagonalizable.

$$\underline{AA}^\top = \underline{PD}_1\underline{P}^\top$$
$$\underline{A}^\top\underline{A} = \underline{QD}_2\underline{Q}^\top$$

Proof. Consider a non-zero eigenvalue of $\underline{A}^\top\underline{A}$, λ. The corresponding normalized eigenvector $\underline{v}$ cannot be the zero vector.

$$\underline{A}^\top\underline{Av} = \lambda\underline{v} \neq \underline{0}$$

Multiply both sides by $\underline{v}^\top$.

$$\begin{aligned}\underline{v}^\top\underline{A}^\top\underline{Av} = \underline{v}^\top\lambda\underline{v} &= \lambda\underline{v}^\top\underline{v} \\ &= \lambda\underline{v}\cdot\underline{v}\end{aligned}$$

The left hand side is really

$$(\underline{Av})^\top(\underline{Av}) = (\underline{Av})\cdot(\underline{Av}) = \|\underline{Av}\|^2 \geq 0$$

Since $\underline{v}\cdot\underline{v} = 1$, if λ is non-zero, it must be positive. □

If we order the r non-zero eigenvalues of $\underline{AA}^\top$ in descending order

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$$

The square root of these r non-zero eigenvalues are called the singular values of $\underline{A}$, $\sigma_i = \sqrt{\lambda_i}$, $i = 1, \ldots, r$.

The singular value decomposition of the $m \times n$ real matrix $\underline{A}$ is

$$A = \underline{P\Sigma Q}^\top,$$

where $\underline{\Sigma}$ is $m \times n$ matrix with the r singular values of $\underline{A}$ on its diagonal, and is padded with 0s such that the matrix multiplication makes sense.

$\underline{P}$ is then a $m \times m$ matrix. $\underline{Q}$ is a $n \times n$ matrix. Both $\underline{P}$ and $\underline{Q}$ are orthogonal matrices.

$\underline{A} = \underline{P\Sigma Q}^\top$ is **not unique** in general (Σ is unique).

We can find the orthogonal matrices $\underline{P}$ and $\underline{Q}$ by orthogonally diagonalizing $\underline{AA}^\top$ and $\underline{A}^\top\underline{A}$. But

$$\underline{A} = \underline{P\Sigma Q}^\top \implies \underline{AQ} = \underline{P\Sigma}$$

The i^{th} column of $\underline{AQ}$ is equal to $\underline{Aq}_i$, and the i^{th} column of $\underline{P\Sigma}$ is $\sigma_i\underline{p}_i$. Since σ_i is non-zero, we have

$$\underline{p}_i = \frac{1}{\sigma_i}\underline{Aq}_i$$

Similarly,

$$\underline{q}_i = \frac{1}{\sigma_i}\underline{A}^\top\underline{p}_i$$

And by definition,

$$\underline{A}^\top\underline{Aq}_i = \sigma_i^2\underline{q}_i \qquad \underline{AA}^\top\underline{p}_i = \sigma_i^2\underline{p}_i$$

When $m \times n$ matrix $\underline{A}$ is tall ($m > n$), diagonalize $\underline{A}^\top\underline{A}$ first (this gives us the rows for $\underline{Q}$, which is a smaller square matrix than $\underline{P}$).

When $n < m$, then diagonalize $\underline{AA}^\top$.

Let $\underline{A} = \underline{P\Sigma Q}^\top$. $\underline{A}$ has r non-zero eigenvalues. Then

- $N(\underline{A}) = \text{span}\left\{\underline{q}_{r+1}, \ldots, \underline{q}_n\right\}$
- $R(\underline{A}) = \text{span}\left\{\underline{p}_1, \ldots, \underline{p}_r\right\}$

These two facts imply:

- rank($\underline{A}$) is the number of nonzero singular values
- $\left[R(\underline{A})\right]^\perp = N(\underline{A}^\top) = \text{span}\left\{\underline{p}_{r+1}, \ldots, \underline{p}_m\right\}$
- $\|\underline{A}\|$ is equal to the largest singular value
- $\|\underline{A}^{-1}\|$ is equal to inverse of the smallest singular value
- cond($\underline{A}$) is equal to σ_1/σ_r
- The Hilbert-Schmidt norm:

$$\|\underline{A}\|_{HS}^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2$$

Proof. We will show that the null space of $\underline{A}$ is spanned by $\left\{\underline{q}_{r+1}, \ldots, \underline{q}_n\right\}$. Let's consider x made of a linear combination of these vectors.
Then,

$$\begin{aligned}\underline{Ax} = \sum_{i=r+1}^{n} w_i\underline{Aq}_i &= \sum_{i=r+1}^{n} w_i\underline{P\Sigma Q}^\top\underline{q}_i \\ &= \sum_{i=r+1}^{n} w_i\underline{P\Sigma}e_i\end{aligned}$$

Σ has non-zero entries on the diagonal up to column r. So over $i \in [r+1, n]$, $\underline{\Sigma}e_i = 0$ □

Proof. We will show that the range of $\underline{A}$ is spanned by $\left\{\underline{p}_1, \ldots, \underline{p}_r\right\}$. By definition,

$$R(\underline{A}) = \left\{\underline{Ax} | x \in \mathbb{R}^n\right\}$$

Then,

$$\underline{Ax} = \underline{P\Sigma}\underbrace{\underline{Q}^\top\underline{x}}_{\underline{y}}$$

The i^{th} entry of $\underline{\Sigma y}$ is $\sigma_i y_i$ if $i \leq r$, and 0 if $i > r+1$. So

$$R(\underline{A}) = \left\{\sum_{j=1}^{r}\sigma_j y_j\underline{p}_j \,|\, \underline{y} \in \mathbb{R}^n\right\}$$

$\sigma_j y_j$ are scalars. □

Proof. We will show that operator norm of $\underline{A}$ is equal to the largest singular value of $\underline{A}$. By definition, $\|\underline{A}\| = \max_{\|x\|=1}\|\underline{Ax}\|$. Then,

$$\max_{\|x\|=1}\|\underline{P\Sigma Q}^\top\underline{x}\|$$

Since both $\underline{P}$ and $\underline{Q}$ are orthogonal matrices, they have norm 1. Then

$$\max_{\|x\|=1}\|\underline{\Sigma x}\|$$

Clearly, the maximum stretch would be equal to the largest singular value. □

Principal component analysis

Consider a $n \times p$ matrix $\underline{X}$. Let each row, $\underline{x}_i$, $i = 1, 2 \ldots, n$ of $\underline{X}$ represent a data point from $\mathbb{R}^p$.

- Assume the data is normalised, the mean of each column of $\underline{X}$ is 0.

We define k^{th} weight vector of $\underline{X}$ to be the the unit vector $\underline{w}_k$ that maximizes

$$\|\underline{X}_k\underline{w}_k\|^2,$$

where $\underline{X}_k$ is the projection of $\underline{X}$ onto the orthogonal compliment of the span of the $1, \ldots, k-1$ weight vectors.

$$\underline{X}_k = \underline{X} - \sum_{j=1}^{k-1}\underline{X}w_j\underline{w}_j^\top$$

The projection coefficient $\underline{x}_i\cdot\underline{w}_k$ is the k^{th} principal component of $\underline{x}_i$.

- The first weight vectors points the direction of maximum variance in the data
- The second weight vector is orthogonal to the first

The weight vectors $\underline{w}_1, \ldots, \underline{w}_p$ are the columns of $\underline{Q}$, $\underline{q}_1, \ldots, \underline{q}_n$, corresponding to singular values $\sigma_1 > \cdots > \sigma_p > 0$.

Pseudoinverse, least squares, and the SVD expansion

Recall in LSA, we solved the normal equation:

$$\underline{A}^\top \underline{A}\underline{z} = \underline{A}^\top \underline{b} \implies \underline{z} = \left(\underline{A}^\top \underline{A}\right)^{-1} \underline{A}^\top \underline{b}$$

By definition, the **left-inverse** of $\underline{A}$ is a matrix $\underline{A}^+$ that satisfies

$$\underline{A}^+ \underline{A} = \underline{I}$$

Let $\underline{A} \in \mathbb{R}^{m \times n}$, where $\text{rank}(\underline{A}) = n$ (so there are no zero singular values) and $m \geq n$. $\underline{A}$ has the SVD equal to $\underline{P\Sigma Q}^\top$.

The left-pseudoinverse of $\underline{A}$ is given by

$$\underline{A}^+ = \underline{Q\Sigma}^+ \underline{P}^\top,$$

where $\underline{\Sigma}^+$ is the transpose of $\underline{\Sigma}$ with all of its singular values inverted.

- If $\underline{A}$ is invertible, then $\underline{A}^+ = \underline{A}^{-1}$

Let $\underline{A}$ be a $m \times n$ matrix and $\underline{b} \in \mathbb{R}^m$. The LSA of the system $\underline{Ax} = \underline{b}$ is given by $\underline{x} = \underline{A}^+ \underline{b}$.

(**SVD expansion**) If $\underline{A} = \underline{P\Sigma Q}^\top$, the SVD expansion of $\underline{A}$ is

$$\underline{A} = \sum_{i=1}^{r} \sigma_i \underline{p}_i \underline{q}_i^\top$$

each $\underline{p}_i \underline{q}_i^\top$ is a $m \times n$ matrix of rank 1.

Inverting noisy systems

Consider a linear system with noise, $\underline{e}$.

$$\underline{Ax} = \underline{b} + \underline{e}$$

Assuming $\underline{e}$ is mostly orthogonal to $\underline{p}_i$ corresponding to large singular values of $\underline{A}$ (in other words, $\underline{e}$ is mostly unrelated to $\underline{A}$), then it is better to use a truncated SVD expansion of the pseudoinverse.

(**SVD expansion of the pseudoinverse**)

$$\underline{A}_k^+ = \sum_{i=1}^{k} \frac{1}{\sigma_i} \underline{q}_i \underline{p}_i^\top$$

So that

$$\underline{x} = \underline{A}_k^+ (\underline{b} + \underline{e})$$

since the small singular values of $\underline{A}$ (which become the large singular values of $\underline{A}^{-1}$) are discarded, and we attenuate the contribution of $\underline{A}_k^+ \underline{e}$.

Power method for computing eigenvalues

Suppose $\underline{A}$ is a square matrix, with an eigenbasis. The **dominant eigenvalue**, λ, of a square matrix is the eigenvalue with multiplicity 1 and the **greatest absolute** value of all eigenvalues.

For a $n \times n$ matrix $\underline{A}$,

- $\lambda_1, \ldots, \lambda_n$ corresponds to $v_1, \ldots, v_n$ eigenvectors
- The eigenvectors form an eigenbasis

We can expand any $\underline{x} \in \mathbb{R}^n$ as a linear combination of the eigenbasis vectors. We define λ_1 to be the dominant eigenvalue. Then

$$\begin{aligned}
\underline{A}^k \underline{x} &= c_1 \underline{A}^k \underline{v}_1 + \cdots + c_n \underline{A}^k \underline{v}_n \\
&= c_1 \lambda_1^k \underline{v}_1 + \cdots + c_n \lambda_n^k \underline{v}_n \\
&= \lambda_1^k \left(c_1 \underline{v}_1 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \underline{v}_n \right)
\end{aligned}$$

For large k,

$$\underline{A}^k \underline{x} = \lambda_1^k \left[c_1 \underline{v}_1 + \underbrace{\cdots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \underline{v}_n}_{\to 0} \right]$$

So we get a scaled version of the dominant eigenvector.

(**Normalized power iteration**) For large k, $\underline{A}^k \underline{x}$ can get very large (or very small, if $\lambda_1 < 1$). The normalized power iteration algorithm prevents overflow or underflow.

1. Choose an arbitrary starting vector $\underline{x}_0$
2. For each $k = 1, 2, \ldots$, $\underline{y}_k = \underline{A}\underline{x}_{k-1}$
3. Set $\underline{x}_k = \underline{y}_k / \|\underline{y}_k\|_\infty$ (or any other norm is also adequate)
4. Repeat

For large k, $\underline{x}_k \to \pm \underline{v}_1$, and $\|y_k\|_\infty \to |\lambda_1|$

(**Inverse iteration**) Finds the dominant eigenvalue of $\underline{A}^{-1}$, which is the inverse of the smallest eigenvalue of $\underline{A}$

1. Choose an arbitrary starting vector $\underline{x}_0$
2. For each $k = 1, 2, \ldots$, solve $\underline{A}\underline{y}_k = \underline{x}_{k-1}$ for $\underline{y}_k$
3. Set $\underline{x}_k = \underline{y}_k / \|\underline{y}_k\|_\infty$
4. Repeat

$\|y_k\|_\infty$ approaches the largest eigenvalue of $\underline{A}^{-1}$, or the inverse of the smallest eigenvalue of $\underline{A}$

Let $s \in \mathbb{R}$ (we can pick anything we like), and $\underline{A}$ is square matrix with real eigenvalues, $\lambda_1, \ldots, \lambda_n$, then one of the two following options hold:

1. s is an eigenvalue of $\underline{A}$. Then by definition,
$$\det\left(\underline{A} - s\underline{I}\right) = 0$$
i.e. $\underline{A}$ is not invertible.
2. An eigenvalue of $(\underline{A} - s\underline{I})^{-1}$ is
$$\frac{1}{\lambda_j - s}$$
and the corresponding eigenvectors are the same as for $\underline{A}$

Proof. We show that the second case holds. Consider an eigenvalue eigenvector pair of $\underline{A}$. Then

$$\underline{Av} = \lambda \underline{v}$$

$(\underline{A} - \underline{I}s)$ acting on $\underline{v}$ is then

$$\begin{aligned}
\underline{Av} - s\underline{Iv} &= \lambda \underline{v} - s\underline{v} \\
(\underline{A} - \underline{I}s)\underline{v} &= (\lambda - s)\underline{v}
\end{aligned}$$

Since $(\underline{A} - \underline{I}s)$ is invertible,

$$(\underline{A} - \underline{I}s)^{-1} \underline{v} = \frac{1}{\lambda - s} \underline{v}$$

□

(**Modified power method**) Finds an eigenvalue of $\underline{A}$, given some real number s such that

$$|\lambda_j - s| < |\lambda_i - s| \qquad \forall i \neq j$$

1. Compute $(\underline{A} - s\underline{I})^{-1}$
2. Applying the normalized power method on $(\underline{A} - s\underline{I})^{-1}$ gets us to the eigenvector of the eigenvalue
$$\frac{1}{\lambda - s}$$
(since $|\lambda_j - s| < |\lambda_i - s| \ \forall i \neq j$)

Markov chains

The **adjacency matrix** of a graph of N vertices is the $N \times N$ matrix $\underline{A}$ where

$$A_{ij} = \begin{cases} 1 & \text{if nodes } i \text{ and } j \text{ are linked} \\ 0 & \text{otherwise} \end{cases}$$

A **stochastic matrix** $\underline{P}$ is a square matrix

$$\underline{P} = \begin{bmatrix} P_{11} & \dots & P_{1i} \\ \vdots & \ddots & \vdots \\ P_{i1} & \dots & P_{ii} \end{bmatrix}$$

with the following properties:

- Each entry in is in $[0, 1]$
- Each column sum up to 1
- The first column represents state 1. The i^{th} entry on column 1 represents the probability of the system to transition from state 1 to state i.

A **state vector** is a column vector

$$\underline{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_i \end{bmatrix}$$

where

- Each entry is between 0 to 1
- The sum of the entries equal 1 ($\|\underline{v}\|_1 = 1$)
- The i^{th} row represents the probability of being at state i during some time

Let $\underline{x}_n$ be the state vector of the system at time n. And $\underline{P}$ be a corresponding stochastic matrix. $\underline{x}_{n+1}$ is given by

$$\underline{x}_{n+1} = \underline{P}\underline{x}_n$$

and in general

$$\underline{x}_n = \underline{P}^n \underline{x}_0$$

Given a Markov chain with transition matrix $\underline{P}$, the **steady-state vector**

$$\underline{x}_\infty = \lim_{n\to\infty} \underline{P}^n \underline{x}_0$$

is the normalized eigenvector of $\underline{P}$ corresponding to $\lambda = 1$. To find the steady-state solution:

1. Solve the system $(\underline{P} - \underline{I})\,\underline{x} = \underline{0}$

(**Properties of stochastic matrices**) Let $\underline{P}$ be a stochastic matrix,

- If $\underline{v}$ is a state vector, then $\underline{P}\underline{v}$ is also a state vector
- $\underline{P}$ has an eigenvalue of $\lambda = 1$
- All other eigenvalues of $\underline{P}$ satisfy $|\lambda_j| \le 1$
- The eigenvector of $\lambda = 1$ has non-negative entries
- If all entries of $\underline{P}$ are > 0, then $\underline{\lambda} = 1$ is the dominant eigenvalue, and all other eigenvalues satisfy $|\lambda_j| < 1$ ($\lambda = 1$ has multiplicity 1, and $\underline{P}$ has a single steady state solution)

When

- its not clear whether there is a dominant eigenvalue (there are zero entries in $\underline{P}$)
- or there are multiple $|\lambda_i| = 1$, there is no steady state

we can apply **damping** to the matrix.

Let $\underline{Q}$ be a stochastic matrix the same size as $\underline{P}$, and all entries of $\underline{Q}$ are the same. Pick $\alpha \in [0, 1]$, and the **google matrix** is

$$\underline{G} = \alpha\underline{P} + (1 - \alpha)\underline{Q}$$

The page rank of all the sites are given by the eigenvector to $\lambda = 1$, **normalized using the 1-norm**.

Complex vector space

Definition:

$$\mathbb{C}^n = \left\{ \underline{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} : z_j \in \mathbb{C},\, j = 1, \dots, n \right\}$$

(**Complex inner product**)

$$\langle \underline{w}, \underline{z} \rangle = \sum_{k=1}^{n} \overline{w}_k z_k$$

We made this change since we want

$$\langle \underline{z}, \underline{z} \rangle = \|\underline{z}\|_2^2$$

By a property of the modulus,

$$|z|^2 = \overline{z}z$$

So

$$\langle \underline{z}, \underline{z} \rangle = \sum_{k=1}^{n} \overline{z}_k z_k = \sum_{k=1}^{n} |z_k|^2$$

- We may have also defined it by having the complex conjugate on the second term. This will change its properties

Under the new inner product:

- Non-commutative: $\langle \underline{w}, \underline{z} \rangle = \overline{\langle \underline{z}, \underline{w} \rangle}$
- Multiplication by a constant:

$$\langle s\underline{w}, \underline{z} \rangle = \overline{s}\langle \underline{w}, \underline{z} \rangle$$
$$\langle \underline{w}, s\underline{z} \rangle = s\langle \underline{w}, \underline{z} \rangle$$

- Matrix product: $\langle \underline{w}, \underline{z} \rangle = \overline{\underline{w}}^{\mathsf{T}} \underline{z}$
- For $\underline{A} \in \mathbb{C}^{m\times n}$,

$$\langle \underline{A}\underline{z}, \underline{w} \rangle = \langle \underline{z}, \overline{\underline{A}}^{\mathsf{T}} \underline{w} \rangle$$

Remember we can also denote the conjugate transpose as $\underline{A}^*$. This is also called the adjoint of $\underline{A}$

Orthonormal basis in $\mathbb{C}^n$

A set vectors $\{\underline{q}_1, \dots, \underline{q}_n\}$ forms an orthonormal basis if

$$\langle \underline{q}_i, \underline{q}_j \rangle = \delta_{ij}$$

Any orthonormal basis for $\mathbb{R}^n$ is also an orthonormal basis for $\mathbb{C}^n$

We can always expand any vector as a linear combination of orthonormal basis vectors. Given $\{\underline{q}_1, \dots, \underline{q}_n\}$, then

$$\underline{v} = \sum_{k=1}^{n} c_k \underline{q}_k = \underbrace{\begin{bmatrix} \underline{q}_1 & \cdots & \underline{q}_n \end{bmatrix}}_{\underline{Q}} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

We can show that the coefficients are given by

$$c_j = \left\langle \underline{q}_j, \underline{v} \right\rangle = \overline{\underline{q}}_j^{\mathsf{T}} \underline{v}$$

Then,

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \overline{\underline{q}}_1^{\mathsf{T}} \underline{v} \\ \vdots \\ \overline{\underline{q}}_n^{\mathsf{T}} \underline{v} \end{bmatrix} = \overline{\underline{Q}}^{\mathsf{T}} \underline{v} = \underline{Q}^* \underline{v}$$

Since

$$\underline{v} = \underline{Q}\underline{c} \qquad\qquad \underline{c} = \underline{Q}^* \underline{v}$$

It means that

$$\underline{Q}\underline{Q}^* = \underline{Q}^*\underline{Q} = \underline{I}$$

It follows that $\underline{Q}$ is an *unitary matrix* (if $\underline{Q}$ had only real entries, then it is an *orthogonal matrix*).

(**Parseval's Identity**) Let $\left\{\underline{q}_j\right\}_{j=1}^{n}$ be an orthonormal basis of some vector space $\mathcal{V}$. Then for all $v \in \mathcal{V}$, we have

$$\sum_{j=1}^{n} \left| \left\langle \underline{q}_j, \underline{v} \right\rangle \right|^2 = \|\underline{v}\|_2^2$$

Proof. We can always expand $\underline{v}$ in terms of the basis vectors. By definition, the two norm squared is equal the inner product of the vector with it self. So

$$\begin{aligned}\langle \underline{v}, \underline{v} \rangle &= \left\langle \sum_{j=1}^{n} c_j \underline{q}_j, \sum_{k=1}^{n} c_k \underline{q}_k \right\rangle \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{c}_j c_k \langle \underline{q}_j, \underline{q}_k \rangle \\ &= \sum_{j=1}^{n} \overline{c}_j c_j \langle \underline{q}_j, \underline{q}_j \rangle \\ &= \sum_{j=1}^{n} |c_j|^2 \\ &= \sum_{j=1}^{n} |\langle \underline{q}_j, \underline{v} \rangle|^2\end{aligned}$$

□

Discrete Fourier transform

Motivation:

- Fourier's theorem tells us that signals can be decomposed by

$$y(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i t n/T}$$

  a weighted superposition of pure oscillations with frequency n/T

- But computing the infinite sum is impractical. So we sample the the signal at discrete points

The N**-th root of unity** is a complex number w such that $w^N = 1$. Let $w_N = \exp(2\pi i/N)$, then

$$1, w_N, w_N^2, \ldots, w_N^{N-1}$$

are all the N-th roots of unity. (Since $w_N^N = 1$ we stop at $N-1$)

Using Euler's formula we can also write w_N^k as

$$w_N^k = \cos\left(\frac{2\pi k}{N}\right) + i \sin\left(\frac{2\pi k}{N}\right)$$

Let $w_N = \exp(2\pi i/N)$. Given an integer $k \neq 0$ or some integer multiple of N, then

$$\sum_{n=0}^{N-1} w_N^{nk} = 0$$

Proof. We draw on the geometric series formula:

$$\sum_{n=0}^{N-1} r^n = \frac{1-r^N}{1-r}$$

which holds for $r \neq 1$. This is true for $r = w_N^k$ since $k \neq 0$ and k is not an integer multiple of N. So

$$\sum_{n=0}^{N-1} w_N^{nk} = \frac{1-w_N^{kN}}{1-w_N^k} = \frac{1-1}{1-w_N^k} = 0$$

□

The **Fourier basis** in $\mathbb{C}^n$ is the set of vectors $\underline{f}_0, \ldots, \underline{f}_{N-1}$ such that

$$\underline{f}_k = \begin{bmatrix} 1 & w_N^k & w_N^{2k} & \ldots & w_N^{(N-1)k} \end{bmatrix}^\intercal$$

The Fourier basis is an orthogonal basis of $\mathbb{C}^n$. Such that

$$\left\langle \underline{f}_k, \underline{f}_\ell \right\rangle = N \delta_{k\ell}$$

Proof. By definition,

$$\left\langle \underline{f}_k, \underline{f}_\ell \right\rangle = \sum_{n=0}^{N-1} \overline{w}_N^{nk} w_N^{n\ell} = \sum_{n=0}^{N-1} w_N^{n(\ell-k)}$$

Clearly, if $\ell = k$, then each $w_N^{n0} = 1$, and the sum is N. Otherwise, the sum is zero we showed using the finite geometric series formula. □

We can expand any vector in $\mathbb{C}^n$ in terms of a linear combination of the Fourier basis.

$$\underline{x} = \sum_{n=0}^{N-1} \frac{\left\langle \underline{f}_n, \underline{x} \right\rangle}{\left\langle \underline{f}_n, \underline{f}_n \right\rangle} \underline{f}_n = \frac{1}{N} \sum_{n=0}^{N-1} \left\langle \underline{f}_n, \underline{x} \right\rangle \underline{f}_n$$

We can write the expansion as a matrix-vector product:

$$\underline{x} = \frac{1}{N} \begin{bmatrix} \underline{f}_0 \cdots \underline{f}_{N-1} \end{bmatrix} \begin{bmatrix} \left\langle \underline{f}_0, \underline{x} \right\rangle \\ \vdots \\ \left\langle \underline{f}_{N-1}, \underline{x} \right\rangle \end{bmatrix}$$

For each inner product, we can write it as

$$\left\langle \underline{f}_n, \underline{x} \right\rangle = \overline{\underline{f}}_n^\intercal \underline{x}$$

So

$$\underline{x} = \frac{1}{N} \begin{bmatrix} \underline{f}_0 \cdots \underline{f}_{N-1} \end{bmatrix} \begin{bmatrix} \overline{\underline{f}}_0^\intercal \\ \vdots \\ \overline{\underline{f}}_{N-1}^\intercal \end{bmatrix} \underline{x}$$

We define the Fourier matrix as

$$\underline{F}_N = \begin{bmatrix} \overline{\underline{f}}_0^\intercal \\ \overline{\underline{f}}_1^\intercal \\ \vdots \\ \overline{\underline{f}}_{N-1}^\intercal \end{bmatrix} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \overline{w}_N & \ldots & \overline{w}_N^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \overline{w}_N^{N-1} & \ldots & \overline{w}_N^{(N-1)(N-1)} \end{bmatrix}$$

such that

$$\text{DFT}(\underline{x}) = \underline{F}_N \underline{x}$$

Properties of the DFT:

1. Let $x \in \mathbb{R}^n$, and $\underline{y} = \text{DFT}(\underline{x}) = \underline{F}_N \underline{x}$, then $\overline{y}_k = y_{N-k}$
2. For each $k = 0, \ldots, N-1$, $\text{DFT}(\underline{f}_k) = N \underline{e}_k$ (i.e., $\text{DFT}(\underline{f}_0) = N \underline{e}_0$, where $\underline{e}_0$ is such that its first entry is 1, and the rest are 0)
3. $\underline{F}_n / \sqrt{N}$ is unitary

The inverse DFT:

$$\text{INVDFT}(\underline{y}) = \frac{1}{N} \underline{F}_N^* \underline{y}$$

- Some sources define the DFT as the one with the scaling factor. But this is a matter of preference. If we only want to compute DFT but not INVDFT, then we can use DFT without the scaling factor so we don't have to worry about scaling

Frequency, amplitude and phase

The DFT computes the amplitude and phase of each frequency in an input signal.

A signal of length N is a vector $x \in \mathbb{C}^N$

$$\underline{x} = \begin{bmatrix} x_0 & x_1 & \ldots & x_{N-1} \end{bmatrix}$$

Consider a sinusoidal signal

$$\underline{x}(\underline{t}) = A \cos(2\pi k \underline{t} + \phi)$$

where

$$\underline{n} = \begin{bmatrix} 0 & 1 & 2 & \ldots & N-1 \end{bmatrix}^\intercal$$

and $\underline{t} = \underline{n}/N$ We define k to be the frequency, A the amplitude, and ϕ the phase. Writing everything out,

$$\underline{A} \cos(2\pi k \underline{t} + \phi) = \begin{bmatrix} A\cos(\phi) \\ A\cos(2\pi k(1/N) + \phi) \\ \vdots \\ A\cos(2\pi k(N-1)/N + \phi) \end{bmatrix}$$

The Fourier basis vectors can be written as:

$$\underline{f}_k = \cos(2\pi k \underline{t}) + i \sin(2\pi k \underline{t})$$

$$\frac{\underline{f}_k + \overline{\underline{f}}_k}{2} = \cos(2\pi k \underline{t})$$

$$\frac{\underline{f}_k - \overline{\underline{f}}_k}{2i} = \sin(2\pi k \underline{t})$$

And

$$\overline{\underline{f}}_k = \underline{f}_{N-k}$$

Let $\underline{x} = A\cos(2\pi k \underline{t} + \phi)$, then

$$\text{DFT}(x) = \frac{AN}{2} e^{i\phi} \underline{e}_k + \frac{AN}{2} e^{-i\phi} \underline{e}_{N-k}$$

Remember k is indexed starting from 0.