1 MATH FACTS

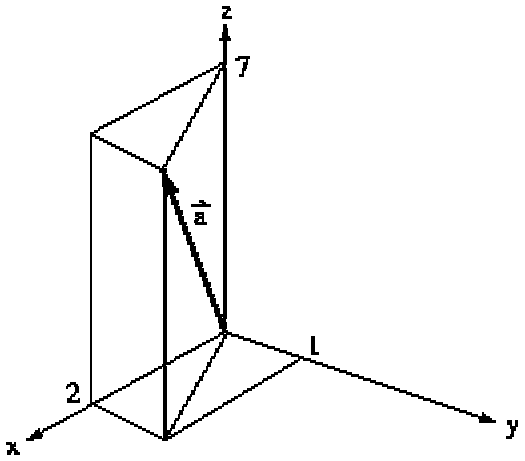
1.1 Vectors

1.1.1 Definition

We use the overhead arrow to denote a column vector, i.e., a *number with a direction*. For example, in three-space, we write

$$\vec{a} = \begin{Bmatrix} 2 \\ 1 \\ 7 \end{Bmatrix}.$$

The elements of a vector have a graphical interpretation, which is particularly easy to see in two or three dimensions.



1. Vector addition is pointwise.

$$\begin{Bmatrix} 2 \\ 1 \\ 7 \end{Bmatrix} + \begin{Bmatrix} 3 \\ 3 \\ 2 \end{Bmatrix} = \begin{Bmatrix} 5 \\ 4 \\ 9 \end{Bmatrix}. \qquad \vec{a} + \vec{b} = \vec{c}$$

   Graphically, addition is stringing the vectors together head to tail.

2. Scalar multiplication is pointwise.

$$-2 \times \begin{Bmatrix} 2 \\ 1 \\ 7 \end{Bmatrix} = \begin{Bmatrix} -4 \\ -2 \\ -14 \end{Bmatrix}.$$

1.1.2 Vector Magnitude

The total length of a vector of dimension m, its Euclidean norm, is given by

$$\|\vec{x}\| = \sqrt{\sum_{i=1}^{m} x_i^2};$$

this scalar is commonly used to normalize a vector to length one.

1.1.3 Vector Dot Product

The dot product of two vectors is the sum of the products of the elements:

$$\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y} = \sum_{i=1}^{m} x_i y_i.$$

The dot product also satisfies

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \|\vec{y}\| \cos\theta,$$

where θ is the angle between the vectors.

1.1.4 Vector Cross Product

The cross product of two three-dimensional vectors is another vector, $\vec{x} \times \vec{y} = \vec{z}$, whose

1. direction is normal to the plane formed by the two vectors,
2. direction is given by the right-hand rule, rotating from $\vec{x}$ to $\vec{y}$,
3. magnitude is the area of the parallelogram formed by the two vectors - the cross product of two parallel vectors is zero - and
4. (signed) magnitude is equal to $\|\vec{x}\| \|\vec{y}\| \sin\theta$, where θ is the angle between the two vectors, measured from $\vec{x}$ to $\vec{y}$.

The schoolbook formula is

$$\vec{x} \times \vec{y} = \begin{Bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{Bmatrix}.$$

1.2 Matrices

1.2.1 Definition

A matrix, or array, is equivalent to a set of row vectors, arranged side by side, say

$$A = [\vec{a} \ \vec{b}] = \begin{bmatrix} 2 & 3 \\ 1 & 3 \\ 7 & 2 \end{bmatrix}.$$

This matrix has three rows ($m = 3$) and two columns ($n = 2$); a vector is a special case of a matrix with one column. Matrices, like vectors, permit pointwise addition and scalar multiplication. We usually use an upper-case symbol to denote a matrix.

1.2.2 Multiplying a Vector by a Matrix

If A_{ij} denotes the element of matrix A in the i'th row and the j'th column, then the multiplication $\vec{c} = A\vec{v}$ is constructed as:

$$c_i = A_{i1} v_1 + A_{i2} v_2 + \cdots + A_{in} v_n = \sum_{j=1}^{n} A_{ij} v_j,$$

where n is the number of columns in A. $\vec{c}$ will have as many columns as A has rows (m). Note that this multiplication is well-defined only if $\vec{v}$ has as many rows as A has columns; they have consistent *inner dimension* n. The product $\vec{v}A$ would be well-posed only if A had one row, and the proper number of columns. There is another important interpretation of this vector multiplication: Let the subscript $:$ indicate all rows, so that each $A_{:j}$ is the j'th column vector. Then

$$\vec{c} = A\vec{v} = A_{:1} v_1 + A_{:2} v_2 + \cdots + A_{:n} v_n.$$

We are multiplying column vectors of A by the scalar elements of $\vec{v}$.

1.2.3 Multiplying a Matrix by a Matrix

The multiplication $C = AB$ is equivalent to a side-by-side arrangement of column vectors $C_{:j} = AB_{:j}$, so that

$$C = AB = [AB_{:1} \ AB_{:2} \ \cdots \ AB_{:k}],$$

where k is the number of columns in matrix B. The same inner dimension condition applies as noted above: the number of columns in A must equal the number of rows in B. Matrix multiplication is:

1. Associative. $(AB)C = A(BC)$.
2. Distributive. $A(B+C) = AB + AC$, $(B+C)A = BA + CA$.
3. NOT Commutative. $AB \neq BA$, except in special cases.

1.2.4 Common Matrices

Identity

. The identity matrix is usually denoted I, and comprises a square matrix with ones on the diagonal, and zeros elsewhere, e.g.,

$$I_{3x3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The identity always satisfies $AI_{nxn} = I_{mxm}A = A$.

Diagonal Matrices

. A diagonal matrix is square, and has all zeros off the diagonal. For instance, the following is a diagonal matrix:

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

The product of a diagonal matrix with another diagonal matrix is diagonal, and in this case the operation is commutative.

1.2.5 Transpose

The transpose of a vector or matrix, indicated by a T superscript results from simply swapping the row-column indices of each entry; it is equivalent to ``flipping'' the vector or matrix around the diagonal line. For example,

$$\vec{a} = \begin{Bmatrix} 1 \\ 2 \\ 3 \end{Bmatrix} \longrightarrow \vec{a}^T = \begin{Bmatrix} 1 & 2 & 3 \end{Bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 8 & 9 \end{bmatrix} \longrightarrow A^T = \begin{bmatrix} 1 & 4 & 8 \\ 2 & 5 & 9 \end{bmatrix}.$$

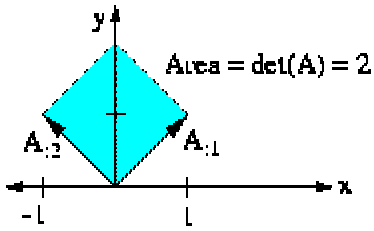
A very useful property of the transpose is

$$(AB)^T = B^T A^T.$$

1.2.6 Determinant

The determinant of a square matrix A is a scalar equal to *the volume* of the parallelepiped enclosed by the constituent vectors. The two-dimensional case is particularly easy to remember, and illustrates the principle of volume:

$$\begin{aligned} det(A) &= A_{11}A_{22} - A_{21}A_{12} \\ det\left(\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right) &= 1 + 1 = 2. \end{aligned}$$



In higher dimensions, the determinant is more complicated to compute. The general formula allows one to pick a row k, perhaps the one containing the most zeros, and apply

$$det(A) = \sum_{j=1}^{j=n} A_{kj}(-1)^{k+j}\Delta_{kj},$$

where Δ_{kj} is the determinant of the sub-matrix formed by neglecting the k'th row and the j'th column. The formula is symmetric, in the sense that one could also target the k'th column:

$$det(A) = \sum_{j=1}^{j=n} A_{jk}(-1)^{k+j}\Delta_{jk}.$$

If the determinant of a matrix is zero, then the matrix is said to be singular - there is no volume, and this results from the fact that the constituent vectors do not span the matrix dimension. For instance, in two dimensions, a singular matrix has the vectors colinear; in three dimensions, a singular matrix has all its vectors lying in a (two-dimensional) plane. Note also that $det(A) = det(A^T)$. If $det(A) \neq 0$, then the matrix is said to be nonsingular.

1.2.7 Inverse

The inverse of a square matrix A, denoted A^{-1}, satisfies $AA^{-1} = A^{-1}A = I$. Its computation requires the determinant above, and the following definition of the $n \times n$ *adjoint* matrix:

$$adj(A) = \begin{bmatrix} (-1)^{1+1}\Delta_{11} & \cdots & (-1)^{1+n}\Delta_{1n} \\ \cdots & \cdots & \cdots \\ (-1)^{n+1}\Delta_{n1} & \cdots & (-1)^{n+n}\Delta_{nn} \end{bmatrix}^T.$$

Once this computation is made, the inverse follows from

$$A^{-1} = \frac{adj(A)}{det(A)}.$$

If A is singular, i.e., $det(A) = 0$, then the inverse does not exist. The inverse finds common application in solving systems of linear equations such as

$$A\vec{x} = \vec{b} \longrightarrow \vec{x} = A^{-1}\vec{b}.$$

1.2.8 Trace

The trace of a matrix is simply the sum of the diagonals:

$$tr(A) = \sum_{i=1}^{n} A_{ii}.$$

1.2.9 Eigenvalues and Eigenvectors

A typical eigenvalue problem is stated as

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

where A is an $n \times n$ matrix, $\vec{x}$ is a column vector with n elements, and λ is a scalar. We ask for what nonzero vectors $\vec{x}$ (right eigenvectors), and scalars λ (eigenvalues) will the equation be satisfied.

Since the above is equivalent to $(A - \lambda I)\vec{x} = \vec{0}$, it is clear that $det(A - \lambda I) = 0$. This observation leads to the solutions for λ; here is an example for the two-dimensional case:

$$
\begin{aligned}
A &= \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \longrightarrow \\
A - \lambda I &= \begin{bmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{bmatrix} \longrightarrow \\
det(A - \lambda I) &= (4-\lambda)(-3-\lambda) + 10 \\
&= \lambda^2 - \lambda - 2 \\
&= (\lambda + 1)(\lambda - 2).
\end{aligned}
$$

Thus, A has two eigenvalues, $\lambda_1 = -1$ and $\lambda_2 = 2$. Each is associated with a *right eigenvector* $\vec{x}$. In this example,

$$
\begin{aligned}
(A - \lambda_1 I)\vec{x}_1 &= \vec{0} \longrightarrow \\
\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix}\vec{x}_1 &= \vec{0} \longrightarrow \\
\vec{x}_1 &= \left\{\sqrt{2}/2,\ \sqrt{2}/2\right\}^T. \\
(A - \lambda_2 I)\vec{x}_2 &= \vec{0} \longrightarrow \\
\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix}\vec{x}_2 &= \vec{0} \longrightarrow \\
\vec{x}_2 &= \left\{5\sqrt{29}/29,\ 2\sqrt{29}/29\right\}^T.
\end{aligned}
$$

Eigenvectors have arbitrary magnitude and sign; they are often normalized to have unity magnitude, and positive first element (as above). A set of n eigenvectors is always linearly independent. The condition that $rank(A - \lambda_i I) = rank(A) - 1$ indicates that there is only one eigenvector for the eigenvalue λ_i. If the left-hand side is less than this, then there are multiple unique eigenvectors that go with λ_i.

The above discussion relates only the right eigenvectors, generated from the equation $A\vec{x} = \lambda\vec{x}$. Left eigenvectors, also useful for many problems, pertain to the transpose of A: $A^T\vec{y} = \lambda\vec{y}$. A and A^T share the same eigenvalues λ, since they share the same determinant. Example:

$$
\begin{aligned}
(A^T - \lambda_1 I)\vec{y}_1 &= \vec{0} \longrightarrow \\
\begin{bmatrix} 5 & 2 \\ -5 & -2 \end{bmatrix}\vec{y}_1 &= \vec{0} \longrightarrow \\
\vec{y}_1 &= \left\{2\sqrt{29}/29,\ -5\sqrt{29}/29\right\}^T. \\
(A^T - \lambda_2 I)\vec{y}_2 &= \vec{0} \longrightarrow \\
\begin{bmatrix} 2 & 2 \\ -5 & -5 \end{bmatrix}\vec{y}_2 &= \vec{0} \longrightarrow \\
\vec{y}_2 &= \left\{\sqrt{2}/2,\ -\sqrt{2}/2\right\}^T.
\end{aligned}
$$

1.2.10 Modal Decomposition

The right and left eigenvectors of a particular eigenvalue have unity dot product, that is $\vec{x}_i^T\vec{y}_i = 1$, with the normalization noted above. The dot product of a left eigenvector with the right eigenvector of *a different eigenvalue* is zero. Thus, if

$$
\begin{aligned}
X &= [\vec{x}_1 \cdots \vec{x}_n], \text{ and} \\
Y &= [\vec{y}_1 \cdots \vec{y}_n],
\end{aligned}
$$

then we have

$$
\begin{aligned}
Y^T X &= I, \text{ or} \\
Y^T &= X^{-1}.
\end{aligned}
$$

Next, construct a diagonal matrix of eigenvalues:

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \cdot & \\ 0 & & \lambda_n \end{bmatrix};$$

it follows that

$$\begin{aligned} AV &= V\Lambda \longrightarrow \\ A &= V\Lambda W^T \\ &= \sum_{i=1}^{n} \lambda_i \vec{v}_i \vec{w}_i^T. \end{aligned}$$

Hence Acan be written as a sum of modal components.[1]

1.2.11 Singular Value

Let $G(s)$be an $m \times n$, possibly complex matrix. The singular value decomposition (SVD) computes three matrices satisfying

$$G = U\Sigma V^*,$$

where $U$is $m \times m$, $\Sigma$is $m \times n$, and $V$is $n \times n$. The star notation indicates a complex-conjugate transpose. The matrix $\Sigma$is diagonal, with the form

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 \\ 0 & 0 & \sigma_p & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $p = min(m, n)$. Each nonzero entry on the diagonal is a real, positive singular value, ordered such that $\sigma_1 > \sigma_2 > \cdots \sigma_p$. The notation is common that $\sigma_1 = \overline{\sigma}$, the maximum singular value, and $\sigma_p = \underline{\sigma}$, the minimum singular value. The auxiliary matrices $U$and $V$are unitary, i.e., they satisfy $X^* = X^{-1}$. Like eigenvalues, the singular values of $G$are related to projections. σ_i represents the Euclidean size of the matrix Galong the i'th singular vector:

$$\begin{aligned} \overline{\sigma} &= max_{\|x\|=1} \|Gx\| \\ \underline{\sigma} &= min_{\|x\|=1} \|Gx\|. \end{aligned}$$

Other properties of the singular value include:

$$\begin{aligned} &\overline{\sigma}(AB) \leq \overline{\sigma}(A)\overline{\sigma}(B) \\ &\overline{\sigma}(A) = \sqrt{\lambda_{max}(A^*A)} \\ &\underline{\sigma}(A) = \sqrt{\lambda_{min}(A^*A)} \\ &\underline{\sigma}(A) = 1/\overline{\sigma}(A^{-1}) \\ &\overline{\sigma}(A) = 1/\underline{\sigma}(A^{-1}) \end{aligned}$$

1.3 Laplace Transform

1.3.1 Definition

The Laplace transform converts time-domain signals into a frequency-domain equivalent. The signal $y(t)$has transform $Y(s)$defined as follows:

$$Y(s) = L(y(t)) = \int_0^\infty y(\tau)e^{-s\tau}d\tau,$$

where $s$is an unspecified complex number; $Y(s)$is considered to be complex as a result. Note that the Laplace transform is linear, and so it is is distributive: $L(x(t) + y(t)) = L(x(t)) + L(y(t))$will hold throughout. The following table gives a list of some useful transform pairs and other properties, for reference.

$$
\begin{aligned}
y(t) &\longleftrightarrow Y(s) \\
\text{(Impulse)} \quad \delta(t) &\longleftrightarrow 1 \\
\text{(Unit Step)} \quad 1(t) &\longleftrightarrow \frac{1}{s} \\
\text{(Unit Ramp)} \quad t &\longleftrightarrow \frac{1}{s^2} \\
e^{-at} &\longleftrightarrow \frac{1}{s+a} \\
\sin \omega t &\longleftrightarrow \frac{\omega}{s^2+\omega^2} \\
\cos \omega t &\longleftrightarrow \frac{s}{s^2+\omega^2} \\
e^{-at} \sin \omega t &\longleftrightarrow \frac{\omega}{(s+a)^2+\omega^2} \\
e^{-at} \cos \omega t &\longleftrightarrow \frac{s+a}{(s+a)^2+\omega^2} \\
\frac{1}{b-a}(e^{-at}-e^{-bt}) &\longleftrightarrow \frac{1}{(s+a)(s+b)} \\
\frac{1}{ab}\left[1+\frac{1}{a-b}(be^{-at}-ae^{-bt})\right] &\longleftrightarrow \frac{1}{s(s+a)(s+b)} \\
\frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta \omega_n t} \sin \omega_n \sqrt{1-\zeta^2} t &\longleftrightarrow \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \\
1-\frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta \omega_n t} \sin\left(\omega_n \sqrt{1-\zeta^2} t+\phi\right) &\longleftrightarrow \frac{\omega_n^2}{s(s^2+2\zeta\omega_n s+\omega_n^2)} \\
\left(\phi = \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta}\right) & \\
\text{(Pure Delay)} \quad y(t-\tau)1(t-\tau) &\longleftrightarrow Y(s)e^{-s\tau}
\end{aligned}
$$

The last two properties are of special importance: for control system design, the differentiation of a signal is equivalent to multiplication of its Laplace transform by s; integration of a signal is equivalent to division by s. The other terms that arise will cancel if $y(0) = 0$, or if $y(0)$ is finite, so they are usually ignored.

1.3.2 Convergence

We note first that the value of s affects the convergence of the integral. For instance, if $y(t) = e^t$, then the integral converges only for $Re(s) > 1$, since the integrand is e^{1-s} in this case. Convergence issues are not a problem for evaluation of the Laplace transform, however, because of *analytic continuation*. This result from complex analysis holds that if two complex functions are equal on some arc (or line) in the complex plane, then they are equivalent everywhere. This fact allows us to always pick a value of s for which the integral above converges, and then by extension infer the existence of the general transform.

1.3.3 Convolution Theorem

One of the main points of the Laplace transform is the ease of dealing with dynamic systems. As with the Fourier transform, the convolution of two signals in the time domain corresponds with the multiplication of signals in the frequency domain. Consider a system whose impulse response is $g(t)$, being driven by an input signal $x(t)$; the output is $y(t) = g(t) * x(t)$. The *Convolution Theorem* is

$$y(t) = \int_0^t g(t-\tau)x(\tau)d\tau \iff Y(s) = G(s)X(s).$$

Here's the proof given by Siebert:

$$\begin{aligned} Y(s) &= \int_0^\infty y(t)e^{-st}dt \\ &= \int_0^\infty \left[\int_0^t g(t-\tau)\,x(\tau)\,d\tau\right] e^{-st}\,dt \\ &= \int_0^\infty \left[\int_0^\infty g(t-\tau)\,1(t-\tau)\,x(\tau)\,d\tau\right] e^{-st}dt \\ &= \int_0^\infty x(\tau)\left[\int_0^\infty g(t-\tau)\,1(t-\tau)\,e^{-st}\,dt\right] d\tau \\ &= \int_0^\infty x(\tau)\,G(s)e^{-s\tau}\,d\tau \\ &= G(s)X(s) \end{aligned}$$

When $g(t)$ is the impulse response of a dynamic system, then $y(t)$ represents the output of this system when it is driven by the external signal $x(t)$.

1.3.4 Solution of Differential Equations by Laplace Transform

The Convolution Theorem allows one to solve (linear time-invariant) differential equations in the following way:

1. Transform the system impulse response $g(t)$ into $G(s)$, and the input signal $x(t)$ into $X(s)$, using the transform pairs.
2. Perform the multiplication in the Laplace domain to find $Y(s)$.
3. Ignoring the effects of pure time delays, break $Y(s)$ into partial fractions with no powers of s greater than 2 in the denominator.
4. Generate the time-domain response from the simple transform pairs. Apply time delay as necessary.

Specific examples of this procedure are given in a later section on transfer functions.

---

[1] By carrying out successive multiplications, it can be shown that A^k has its eigenvalues at λ_i^k, and keeps the same eigenvectors as A.