Linear Algebra Refresher

Andreas Adelmann

Notation

$\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}^n$	real numbers, positive reals, n-tuples of reals
$\mathbb{N}$, $\mathbb{C}$	natural numbers $\{0, 1, 2, \ldots\}$, complex numbers
$(a\,..\,b)$, $[a\,..\,b]$	open interval, closed interval
$\langle \ldots \rangle$	sequence (a list in which order matters)
$h_{i,j}$	row i and column j entry of matrix H
V, W, U	vector spaces
$\vec{v}$, $\vec{0}$, $\vec{0}_V$	vector, zero vector, zero vector of a space V
$\mathcal{P}_n$, $\mathcal{M}_{n\times m}$	space of degree n polynomials, $n\times m$ matrices
$[S]$	span of a set
$\langle B, D\rangle$, $\vec{\beta}, \vec{\delta}$	basis, basis vectors
$\mathcal{E}_n = \langle \vec{e}_1, \ldots, \vec{e}_n\rangle$	standard basis for $\mathbb{R}^n$
h, g	homomorphisms (linear maps)
t, s	transformations (linear maps from a space to itself)
$\mathrm{Rep}_B(\vec{v})$, $\mathrm{Rep}_{B,D}(h)$	representation of a vector, a map
$Z_{n\times m}$ or Z, $I_{n\times n}$ or I	zero matrix, identity matrix
$\lvert T\rvert$	determinant of the matrix

The lectures based on the material and book by Prof. Jim Hefferon.

Gauss's Method

Linear systems

1.1 *Definition* A *linear equation* in the variables $x_1, \ldots, x_n$ has the form $a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = d$ where $d \in \mathbb{R}$ is the *constant*.

An n-tuple $(s_1, s_2, \ldots, s_n) \in \mathbb{R}^n$ is a *solution* of, or *satisfies*, that equation if substituting the numbers $s_1, \ldots, s_n$ for the variables gives a true statement: $a_1s_1 + a_2s_2 + \cdots + a_ns_n = d$. A *system of linear equations*

$$\begin{aligned} a_{1,1}x_1 + \ a_{1,2}x_2 + \cdots + \ a_{1,n}x_n &= \ d_1 \\ a_{2,1}x_1 + \ a_{2,2}x_2 + \cdots + \ a_{2,n}x_n &= \ d_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= d_m \end{aligned}$$

has the solution $(s_1, s_2, \ldots, s_n)$ if that n-tuple is a solution of all of the equations.

Example There are three linear equations in this linear system.

$$\begin{aligned} (1/4)x + \ \ y - \ \ z &= \ 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - \ \ z &= \ 3 \end{aligned}$$

Solving a linear system

Example To find the solution of this system

$$\begin{aligned} (1/4)x + \quad y - \quad z &= \ 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - \quad z &= \ 3 \end{aligned}$$

we transform it to one whose solution is easy.

Solving a linear system

Example To find the solution of this system

$$\begin{aligned} (1/4)x + y - z &= 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - z &= 3 \end{aligned}$$

we transform it to one whose solution is easy. Start by clearing the fraction.

$$\xrightarrow{4\rho_1} \quad \begin{aligned} x + 4y - 4z &= 0 \\ x + 4y + 2z &= 12 \\ 2x - 3y - z &= 3 \end{aligned}$$

Solving a linear system

Example To find the solution of this system

$$\begin{aligned}(1/4)x + y - z &= 0\\ x + 4y + 2z &= 12\\ 2x - 3y - z &= 3\end{aligned}$$

we transform it to one whose solution is easy. Start by clearing the fraction.

$$\xrightarrow{4\rho_1}\quad \begin{aligned}x + 4y - 4z &= 0\\ x + 4y + 2z &= 12\\ 2x - 3y - z &= 3\end{aligned}$$

Next use the first row to act on the rows below, eliminating their x terms.

$$\xrightarrow[-2\rho_1+\rho_3]{-\rho_1+\rho_2}\quad \begin{aligned}x + 4y - 4z &= 0\\ +6z &= 12\\ -11y + 7z &= 3\end{aligned}$$

Then swap to bring a y term to the second row.

$$\xrightarrow{\rho_2 \leftrightarrow \rho_3} \begin{aligned} x + \quad 4y - 4z &= 0 \\ -11y + 7z &= 3 \\ 6z &= 12 \end{aligned}$$

Then swap to bring a y term to the second row.

$$\xrightarrow{\rho_2 \leftrightarrow \rho_3} \quad \begin{aligned} x + \quad 4y - 4z &= 0 \\ -11y + 7z &= 3 \\ 6z &= 12 \end{aligned}$$

Now solve the bottom row: $z = 2$. With that, the shape of the transformed system lets us solve for y by substituting into the second row: $-11y + 7(2) = 3$ shows $y = 1$.

Then swap to bring a y term to the second row.

$$\xrightarrow{\rho_2 \leftrightarrow \rho_3} \quad \begin{aligned} x + \quad 4y - 4z &= 0 \\ -11y + 7z &= 3 \\ 6z &= 12 \end{aligned}$$

Now solve the bottom row: $z = 2$. With that, the shape of the transformed system lets us solve for y by substituting into the second row: $-11y + 7(2) = 3$ shows $y = 1$. The shape also lets us solve for x by substituting into the first row: $x + 4(1) - 4(2) = 0$, so that $x = 4$.

1.7 *Definition*

A matrix that has undergone Gaussian elimination is said to be in row echelon form or, more properly, "reduced echelon form" or "row-reduced echelon form." Such a matrix has the following characteristics:

1) All zero rows are at the bottom of the matrix
2) The leading entry of each nonzero row after the first occurs to the right of the leading entry of the previous row.
3) The leading entry in any nonzero row is 1.
4) All entries in the column above and below a leading 1 are zero.

Another common definition of echelon form only requires zeros below the leading ones, while the above definition also requires them above the leading ones.

Example

$$\begin{array}{rcl} 2x - 3y - z + 2w &=& -2 \\ x + 3z + 1w &=& 6 \\ 2x - 3y - z + 3w &=& -3 \\ y + z - 2w &=& 4 \end{array} \xrightarrow[-\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2} \begin{array}{rcl} 2x - 3y - z + 2w &=& -2 \\ (3/2)y + (7/2)z &=& 7 \\ w &=& -1 \\ y + z - 2w &=& 4 \end{array}$$

Example

$$\begin{array}{rcl} 2x - 3y - z + 2w &=& -2 \\ x + 3z + 1w &=& 6 \\ 2x - 3y - z + 3w &=& -3 \\ y + z - 2w &=& 4 \end{array} \xrightarrow[-\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2} \begin{array}{rcl} 2x - 3y - z + 2w &=& -2 \\ (3/2)y + (7/2)z &=& 7 \\ w &=& -1 \\ y + z - 2w &=& 4 \end{array}$$

$$\xrightarrow{(-2/3)\rho_2+\rho_4} \begin{array}{rcl} 2x - 3y - z + 2w &=& -2 \\ (3/2)y + (7/2)z &=& 7 \\ w &=& -1 \\ -(4/3)z - 2w &=& -2/3 \end{array}$$

Example

$$\begin{aligned}2x-3y-z+2w&=-2\\ x+3z+1w&=6\\ 2x-3y-z+3w&=-3\\ y+z-2w&=4\end{aligned}\quad\xrightarrow[-\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2}\quad\begin{aligned}2x-3y-z+2w&=-2\\ (3/2)y+(7/2)z&=7\\ w&=-1\\ y+z-2w&=4\end{aligned}$$

$$\xrightarrow{(-2/3)\rho_2+\rho_4}\quad\begin{aligned}2x-3y-z+2w&=-2\\ (3/2)y+(7/2)z&=7\\ w&=-1\\ -(4/3)z-2w&=-2/3\end{aligned}$$

$$\xrightarrow{\rho_3\leftrightarrow\rho_4}\quad\begin{aligned}2x-3y-z+2w&=-2\\ (3/2)y+(7/2)z&=7\\ -(4/3)z-2w&=-2/3\\ w&=-1\end{aligned}$$

Example

$$\begin{aligned} 2x - 3y - z + 2w &= -2 \\ x + 3z + 1w &= 6 \\ 2x - 3y - z + 3w &= -3 \\ y + z - 2w &= 4 \end{aligned} \xrightarrow[-\rho_1+\rho_3]{(-1/2)\rho_1+\rho_2} \begin{aligned} 2x - 3y - z + 2w &= -2 \\ (3/2)y + (7/2)z &= 7 \\ w &= -1 \\ y + z - 2w &= 4 \end{aligned}$$

$$\xrightarrow{(-2/3)\rho_2+\rho_4} \begin{aligned} 2x - 3y - z + 2w &= -2 \\ (3/2)y + (7/2)z &= 7 \\ w &= -1 \\ -(4/3)z - 2w &= -2/3 \end{aligned}$$

$$\xrightarrow{\rho_3 \leftrightarrow \rho_4} \begin{aligned} 2x - 3y - z + 2w &= -2 \\ (3/2)y + (7/2)z &= 7 \\ -(4/3)z - 2w &= -2/3 \\ w &= -1 \end{aligned}$$

The fourth equation says $w = -1$. Substituting back into the third equation gives $z = 2$. Then back substitution into the second and first rows gives $y = 0$ and $x = 1$. The unique solution is $(1, 0, 2, -1)$.

Gauss's Method

1.3 *Theorem* If a linear system is changed to another by one of these operations

1) an equation is swapped with another
2) an equation has both sides multiplied by a nonzero constant
3) an equation is replaced by the sum of itself and a multiple of another

then the two systems have the same set of solutions.

1.4 *Definition* The three operations from Theorem 1.3 are the *elementary reduction operations*, or *row operations*, or *Gaussian operations*. They are *swapping*, *multiplying by a scalar* (or *rescaling*), and *row combination*.

Systems without a unique solution

Example This system has no solution.

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + z &= 8 \\ 2x + 3y + 2z &= 13 \end{aligned}$$

On the left the sum of the first two rows equals the third row, while on the right that is not so. So there is no triple of reals that makes all three equations true.

Systems without a unique solution

Example This system has no solution.

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + z &= 8 \\ 2x + 3y + 2z &= 13 \end{aligned}$$

On the left the sum of the first two rows equals the third row, while on the right that is not so. So there is no triple of reals that makes all three equations true.

Gauss' Method makes the inconsistency clear.

$$\xrightarrow[-2\rho_1+\rho_3]{-\rho_1+\rho_2} \begin{aligned} x + y + z &= 6 \\ y &= 2 \\ y &= 1 \end{aligned} \xrightarrow{-\rho_2+\rho_3} \begin{aligned} x + y + z &= 6 \\ y &= 2 \\ 0 &= -1 \end{aligned}$$

Example This system has infinitely many solutions.

$$\begin{array}{rcl} -x - y + 3z = & 3 \\ x \quad + \; z = & 3 \\ 3x - y + 7z = & 15 \end{array} \xrightarrow[3\rho_1+\rho_3]{\rho_1+\rho_2} \begin{array}{rcl} -x - \quad y + \; 3z = & 3 \\ -y + \; 4z = & 6 \\ -4y + 16z = & 24 \end{array}$$

$$\xrightarrow{-4\rho_2+\rho_3} \begin{array}{r} -x - \quad y + 3z = 3 \\ -y + 4z = 6 \\ 0 = 0 \end{array}$$

Taking $z = 0$ gives $(3, -6, 0)$ while taking $z = 1$ gives $(2, -2, 1)$.

Example This system has infinitely many solutions.

$$\begin{array}{rcl} -x - y + 3z = 3 \\ x \quad + z = 3 \\ 3x - y + 7z = 15 \end{array} \xrightarrow[3\rho_1+\rho_3]{\rho_1+\rho_2} \begin{array}{rcl} -x - y + 3z = 3 \\ -y + 4z = 6 \\ -4y + 16z = 24 \end{array}$$

$$\xrightarrow{-4\rho_2+\rho_3} \begin{array}{r} -x - y + 3z = 3 \\ -y + 4z = 6 \\ 0 = 0 \end{array}$$

Taking $z = 0$ gives $(3, -6, 0)$ while taking $z = 1$ gives $(2, -2, 1)$.

Example It is not the '$0 = 0$' that counts. This also has infinitely many solutions.

$$\begin{array}{r} x - y + z = 4 \\ x + y - 2z = -1 \end{array} \xrightarrow{-\rho_1+\rho_2} \begin{array}{r} x - y + z = 4 \\ 2y - 3z = -5 \end{array}$$

Taking $z = 0$ gives the solution $(3/2, -5/2, 0)$. Taking $z = -1$ gives $(1, -4, -1)$.

Elementary Definitions

Matrices and vectors

2.2 *Definition* An $m \times n$ *matrix* is a rectangular array of numbers with m *rows* and n *columns*. Each number in the matrix is an *entry*.

Example This is a 2×3 matrix

$$B = \begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \end{pmatrix}$$

because it has 2 rows and 3 columns. The entry in row 2 and column 1 is $b_{2,1} = 4$.

2.4 *Definition* A *column vector*, often just called a *vector*, is a matrix with a single column. A matrix with a single row is a *row vector*. The entries of a vector are sometimes called *components*. A column or row vector whose components are all zeros is a *zero vector*.

Matrices and vectors

2.2 *Definition* An $m \times n$ *matrix* is a rectangular array of numbers with m *rows* and n *columns*. Each number in the matrix is an *entry*.

Example This is a 2×3 matrix

$$B = \begin{pmatrix} 1 & -2 & 3 \\ 4 & -5 & 6 \end{pmatrix}$$

because it has 2 rows and 3 columns. The entry in row 2 and column 1 is $b_{2,1} = 4$.

2.4 *Definition* A *column vector*, often just called a *vector*, is a matrix with a single column. A matrix with a single row is a *row vector*. The entries of a vector are sometimes called *components*. A column or row vector whose components are all zeros is a *zero vector*.

We denote vectors with an over-arrow

Example This column vector has three components.

$$\vec{v} = \begin{pmatrix} -1 \\ -0.5 \\ 0 \end{pmatrix}$$

Example This row vector has three components

$$\vec{w} = \begin{pmatrix} -1 & -0.5 & 0 \end{pmatrix}$$

Example This is the two-component zero vector.

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Vector operations

2.6 *Definition* The *vector sum* of $\vec{u}$ and $\vec{v}$ is the vector of the sums.

$$\vec{u} + \vec{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

2.7 *Definition* The *scalar multiplication* of the real number r and the vector $\vec{v}$ is the vector of the multiples.

$$r \cdot \vec{v} = r \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} rv_1 \\ \vdots \\ rv_n \end{pmatrix}$$

Example

$$3 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

$$\text{General} = \text{Particular} + \text{Homogeneous}$$

Form of solution sets

Example This system

$$\begin{aligned} x + 2y - z &= 2 \\ 2x - y - 2z + w &= 5 \end{aligned}$$

has solutions of this form.

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 12/5 \\ -1/5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \qquad z, w \in \mathbb{R}$$

Taking $z = w = 0$ shows that the first vector is a particular solution of the system.

3.2 *Definition* A linear equation is *homogeneous* if it has a constant of zero, so that it can be written as $a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$.

Example From the prior system

$$\begin{aligned} x + 2y - z \phantom{{}+w} &= 2 \\ 2x - y - 2z + w &= 5 \end{aligned}$$

we get this associated system of homogeneous equations.

$$\begin{aligned} x + 2y - z \phantom{{}+w} &= 0 \\ 2x - y - 2z + w &= 0 \end{aligned}$$

The same Gauss's Method steps reduce it to echelon form.

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 0 \\ 2 & -1 & -2 & 1 & 0 \end{array}\right) \xrightarrow{-2\rho_1+\rho_2} \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 0 \\ 0 & -5 & 0 & 1 & 0 \end{array}\right)$$

The vector description of the solution set is like the earlier one but the zero vector is a particular solution.

$$\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} z + \begin{pmatrix} -2/5 \\ 1/5 \\ 0 \\ 1 \end{pmatrix} w \mid z, w \in \mathbb{R}\}$$

3.1 *Theorem* Any linear system's solution set has the form

$$\{\vec{p} + c_1\vec{\beta}_1 + \cdots + c_k\vec{\beta}_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

where $\vec{p}$ is any particular solution and where the number of vectors $\vec{\beta}_1, \ldots, \vec{\beta}_k$ equals the number of free variables that the system has after a Gaussian reduction.

3.3 *Corollary* Solution sets of linear systems are either empty, have one element, or have infinitely many elements.

Summary: Kinds of Solution Sets

		number of solutions of the homogeneous system: one	*infinitely many*
particular solution exists?	*yes*	unique solution	infinitely many solutions
	no	no solutions	no solutions

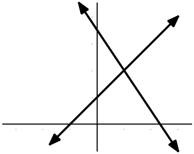
An important special case is when there are the same number of equations as unknowns.

3.4 *Definition* A square matrix is *nonsingular* if it is the matrix of coefficients of a homogeneous system with a unique solution. It is *singular* otherwise, that is, if it is the matrix of coefficients of a homogeneous system with infinitely many solutions.

Geometric Interpretation

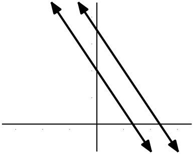
We can draw two-unknown equations as lines. Then the three possibilities for solution sets become clear.

Unique solution



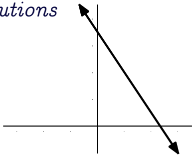
$$\begin{aligned} 3x + 2y &= 7 \\ x - y &= -1 \end{aligned}$$

No solutions



$$\begin{aligned} 3x + 2y &= 7 \\ 3x + 2y &= 4 \end{aligned}$$

Infinitely many solutions



$$\begin{aligned} 3x + 2y &= 7 \\ 6x + 4y &= 14 \end{aligned}$$

This is a nice restatement of the possibilities; the geometry gives us insight into what can happen with linear systems.

Length and angle measures

Length

?? *Definition* The *length* of a vector $\vec{v} \in \mathbb{R}^n$ is the square root of the sum of the squares of its components.

$$|\vec{v}| = \sqrt{v_1^2 + \cdots + v_n^2}$$

Example The length of

$$\begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix}$$

is $\sqrt{1+4+9} = \sqrt{14}$.

For any nonzero vector $\vec{v}$, the length one vector with the same direction is $\vec{v}/|\vec{v}|$. We say that this *normalizes* $\vec{v}$ to unit length.

Dot product

?? *Definition* The *dot product* (or *inner product* or *scalar product*) of two n-component real vectors is the linear combination of their components.

$$\vec{u} \bullet \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Example The dot product of two vectors

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \bullet \begin{pmatrix} 3 \\ -3 \\ 4 \end{pmatrix} = 3 - 3 - 4 = -4$$

is a scalar, not a vector.

The dot product of a vector with itself $\vec{v} \bullet \vec{v} = v_1^2 + \cdots + v_n^2$ is the square of the vector's length.

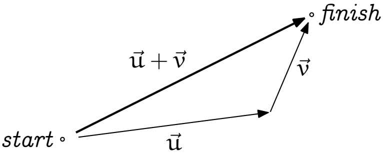
Triangle Inequality

?? *Theorem* For any $\vec{u}, \vec{v} \in \mathbb{R}^n$,

$$|\vec{u} + \vec{v}| \leqslant |\vec{u}| + |\vec{v}|$$

with equality if and only if one of the vectors is a nonnegative scalar multiple of the other one.

This is the source of the familiar saying, “The shortest distance between two points is in a straight line.”



Cauchy-Schwarz Inequality

?? *Corollary* For any $\vec{u}, \vec{v} \in \mathbb{R}^n$,

$$|\vec{u} \bullet \vec{v}| \leqslant |\vec{u}||\vec{v}|$$

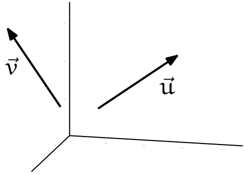
with equality if and only if one vector is a scalar multiple of the other.

Angle measure

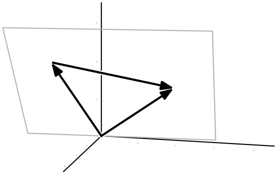
Definition The *angle* between two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ is this.

$$\theta = \arccos\left(\frac{\vec{u} \cdot \vec{v}}{|\vec{u}\,|\,|\vec{v}\,|} \right)$$

We motivate that definition with two vectors in $\mathbb{R}^3$.



If neither is a multiple of the other then they determine a plane, because if we put them in canonical position then the origin and the endpoints make three noncolinear points. Consider the triangle formed by $\vec{u}$, $\vec{v}$, and $\vec{u} - \vec{v}$.



Apply the Law of Cosines: $|\vec{u}-\vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 - 2\,|\vec{u}|\,|\vec{v}|\cos\theta$ where θ is the angle that we want to find. The left side gives

$$\begin{aligned}(u_1-v_1)^2 &+ (u_2-v_2)^2 + (u_3-v_3)^2 \\ &= (u_1^2 - 2u_1v_1 + v_1^2) + (u_2^2 - 2u_2v_2 + v_2^2) + (u_3^2 - 2u_3v_3 + v_3^2)\end{aligned}$$

while the right side gives this.

$$(u_1^2 + u_2^2 + u_3^2) + (v_1^2 + v_2^2 + v_3^2) - 2\,|\vec{u}|\,|\vec{v}|\cos\theta$$

Canceling squares $u_1^2\ \ldots,\ v_3^2$ and dividing by 2 gives the formula.

?? *Corollary* Vectors from $\mathbb{R}^n$ are orthogonal, that is, perpendicular, if and only if their dot product is zero. They are parallel if and only if their dot product equals the product of their lengths.

Vector Spaces Definition and Examples

Vector space

?? *Definition* A *vector space* (over $\mathbb{R}$) consists of a set V along with two operations '+' and '$\cdot$' subject to the conditions that for all vectors $\vec{v}, \vec{w}, \vec{u} \in V$, and all *scalars* $r, s \in \mathbb{R}$:

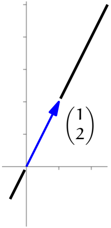
1) the set V is *closed* under vector addition, that is, $\vec{v} + \vec{w} \in V$
2) vector addition is commutative $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
3) vector addition is associative $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$
4) there is a *zero vector* $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$
5) each $\vec{v} \in V$ has an *additive inverse* $\vec{w} \in V$ such that $\vec{w} + \vec{v} = \vec{0}$

Vector space

?? *Definition* A *vector space* (over $\mathbb{R}$) consists of a set V along with two operations '+' and '$\cdot$' subject to the conditions that for all vectors $\vec{v}, \vec{w}, \vec{u} \in V$, and all *scalars* $r, s \in \mathbb{R}$:

1) the set V is *closed* under vector addition, that is, $\vec{v} + \vec{w} \in V$
2) vector addition is commutative $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
3) vector addition is associative $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$
4) there is a *zero vector* $\vec{0} \in V$ such that $\vec{v} + \vec{0} = \vec{v}$ for all $\vec{v} \in V$
5) each $\vec{v} \in V$ has an *additive inverse* $\vec{w} \in V$ such that $\vec{w} + \vec{v} = \vec{0}$
6) the set V is closed under scalar multiplication, that is, $r \cdot \vec{v} \in V$
7) addition of scalars distributes over scalar multiplication
   $(r + s) \cdot \vec{v} = r \cdot \vec{v} + s \cdot \vec{v}$
8) scalar multiplication distributes over vector addition
   $r \cdot (\vec{v} + \vec{w}) = r \cdot \vec{v} + r \cdot \vec{w}$
9) ordinary multipication of scalars associates with scalar multiplication
   $(rs) \cdot \vec{v} = r \cdot (s \cdot \vec{v})$
10) multiplication by the scalar 1 is the identity operation $1 \cdot \vec{v} = \vec{v}$.

Example Let V be the line with slope 2 that passes through the origin in the plane.



$$V = \{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 2x \}$$

It is a set consisting of vectors. Here are some of its infinitely many elements.

$$\begin{pmatrix} 4 \\ 8 \end{pmatrix} \quad \begin{pmatrix} 1/2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} -100 \\ -200 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We will show that this set is a vector space, where the operations are the usual vector addition and scalar multiplication.

Verify conditions (1)-(10) above and arrive at the conclusion: V is a vector space, under the natural addition and scalar multiplication operations.

$\mathbb{R}^n$

The set of n-tall vectors is a vector space under the natural operations. All ten conditions are easy; we will just verify condition (1). Where

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

then the sum

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

is also a member of $\mathbb{R}^n$. (There are no restrictions to check, since every n-tall vector is a member of $\mathbb{R}^n$.)

Example Consider the set of quadratic polynomials.

$$\mathcal{P}_2 = \{ a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R} \}$$

Some members are $3 + 2x + 1x^2$, $10 + 0x + 5x^2$, and $0 + 0x + 0x^2$.

Example Consider the set of quadratic polynomials.

$$\mathcal{P}_2 = \{ a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R} \}$$

Some members are $3 + 2x + 1x^2$, $10 + 0x + 5x^2$, and $0 + 0x + 0x^2$. This is a vector space under the usual operations of polynomial addition

$$(a_0 + a_1 x + a_2 x^2) + (b_0 + b_1 x + b_2 x^2) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

and scalar multiplication.

$$r \cdot (a_0 + a_1 x + a_2 x^2) = (r a_0) + (r a_1)x + (r a_2)x^2$$

Remember the intuition that a vector space is a place where linear combinations can happen. Here is a sample combination in $\mathcal{P}_2$

$$4 \cdot (1 + 2x + 3x^2) - (1/5) \cdot (10 + 5x^2) = 2 + 8x + 11x^2$$

illustrating that a linear combination of quadratic polynomials is a quadratic polynomial.

Subspaces and spanning sets

Example This is not a subspace of $\mathbb{R}^3$.

$$T = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid x + y + z = 1 \}$$

It is a subset of $\mathbb{R}^3$ but it is not a vector space. One condition that it violates is that it is not closed under vector addition: here are two elements of T that sum to a vector that is not an element of T.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

(Another reason that it is not a vector space is that it does not satisfy condition (6). Still another is that it does not contain the zero vector.)

Span

?? *Definition* The *span* (or *linear closure*) of a nonempty subset S of a vector space is the set of all linear combinations of vectors from S.

$$[S] = \{c_1\vec{s}_1 + \cdots + c_n\vec{s}_n \mid c_1, \ldots, c_n \in \mathbb{R} \text{ and } \vec{s}_1, \ldots, \vec{s}_n \in S\}$$

The span of the empty subset of a vector space is its trivial subspace.

No notation for the span is completely standard. The square brackets used here are common but so are 'span(S)' and 'sp(S)'.

Span

?? *Definition* The *span* (or *linear closure*) of a nonempty subset S of a vector space is the set of all linear combinations of vectors from S.

$$[S] = \{c_1\vec{s}_1 + \cdots + c_n\vec{s}_n \mid c_1, \ldots, c_n \in \mathbb{R} \text{ and } \vec{s}_1, \ldots, \vec{s}_n \in S\}$$

The span of the empty subset of a vector space is its trivial subspace.

No notation for the span is completely standard. The square brackets used here are common but so are 'span(S)' and 'sp(S)'.

Example Inside the vector space of all two-wide row vectors, the span of this one-element set

$$S = \{\begin{pmatrix}1 & 2\end{pmatrix}\}$$

is this.

$$[S] = \{\begin{pmatrix}a & 2a\end{pmatrix} \mid a \in \mathbb{R}\} = \{\begin{pmatrix}1 & 2\end{pmatrix}a \mid a \in \mathbb{R}\}$$

Basis

Definition of basis

?? *Definition* A ***basis*** for a vector space is a sequence of vectors that is linearly independent and that spans the space.

Because a basis is a sequence, meaning that bases are different if they contain the same elements but in different orders, we denote it with angle brackets $\langle\vec{\beta}_1, \vec{\beta}_2, \ldots\rangle$.

Definition of basis

?? *Definition* A *basis* for a vector space is a sequence of vectors that is linearly independent and that spans the space.

Because a basis is a sequence, meaning that bases are different if they contain the same elements but in different orders, we denote it with angle brackets $\langle \vec{\beta}_1, \vec{\beta}_2, \ldots \rangle$.

Example This is a basis for $\mathbb{R}^2$.

$$\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

It is linearly independent.

$$c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{aligned} c_1 + c_2 &= 0 \\ -c_1 + c_2 &= 0 \end{aligned} \implies c_1 = 0,\ c_2 = 0$$

And it spans $\mathbb{R}^2$ since

$$c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{aligned} c_1 + c_2 &= x \\ -c_1 + c_2 &= y \end{aligned}$$

has the solution $c_1 = (1/2)x - (1/2)y$ and $c_2 = (1/2)x + (1/2)y$.

Example This is a basis for $\mathcal{M}_{2\times 2}$.

$$\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} \rangle$$

This is another one.

$$\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \rangle$$

Example This is a basis for $\mathbb{R}^3$.

$$\mathcal{E}_3 = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \rangle$$

?? *Definition* For any $\mathbb{R}^n$

$$\mathcal{E}_n = \langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \rangle$$

is the *standard* (or *natural*) basis. We denote these vectors $\vec{e}_1, \ldots, \vec{e}_n$.

?? *Definition* In a vector space with basis B the *representation of $\vec{v}$ with respect to B* is the column vector of the coefficients used to express $\vec{v}$ as a linear combination of the basis vectors:

$$\mathrm{Rep}_B(\vec{v}) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

where $B = \langle \vec{\beta}_1, \ldots, \vec{\beta}_n \rangle$ and $\vec{v} = c_1\vec{\beta}_1 + c_2\vec{\beta}_2 + \cdots + c_n\vec{\beta}_n$. The c's are the *coordinates of $\vec{v}$ with respect to B*.

Example Above we saw that in $\mathcal{P}_1 = \{a + bx \mid a, b \in \mathbb{R}\}$ one basis is $B = \langle 1 + x, 1 - x \rangle$. As part of that we computed the coefficients needed to express a member of $\mathcal{P}_1$ as a combination of basis vectors.

$$a + bx = c_1(1 + x) + c_2(1 - x) \implies c_1 = (a + b)/2,\ c_2 = (a - b)/2$$

For instance, the polynomial $3 + 4x$ has this expression

$$3 + 4x = (7/2) \cdot (1 + x) + (-1/2) \cdot (1 - x)$$

so its representation is this.

$$\mathrm{Rep}_B(3 + 4x) = \begin{pmatrix} 7/2 \\ -1/2 \end{pmatrix}$$

Example With respect to $\mathbb{R}^3$'s standard basis $\mathcal{E}_3$ the vector

$$\vec{v} = \begin{pmatrix} 2 \\ -3 \\ 1/2 \end{pmatrix}$$

has this representation.

$$\mathrm{Rep}_{\mathcal{E}_3}(\vec{v}) = \begin{pmatrix} 2 \\ -3 \\ 1/2 \end{pmatrix}$$

In general, any $\vec{w} \in \mathbb{R}^n$ has $\mathrm{Rep}_{\mathcal{E}_n}(\vec{w}) = \vec{w}$.

Dimension

Definition of dimension

?? *Definition* A vector space is *finite-dimensional* if it has a basis with only finitely many vectors.

Example The space $\mathbb{R}^3$ is finite-dimensional since it has a basis with three elements $\mathcal{E}_3$.

Example The space of quadratic polynomials $\mathcal{P}_2$ has at least one basis with finitely many elements, $\langle 1, x, x^2 \rangle$, so it is finite-dimensional.

Example The space $\mathcal{M}_{2\times2}$ of 2×2 matrices is finite-dimensional. Here is one basis with finitely many members.

$$\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rangle$$

Definition of dimension

?? *Definition* A vector space is *finite-dimensional* if it has a basis with only finitely many vectors.

Example The space $\mathbb{R}^3$ is finite-dimensional since it has a basis with three elements $\mathcal{E}_3$.

Example The space of quadratic polynomials $\mathcal{P}_2$ has at least one basis with finitely many elements, $\langle 1, x, x^2 \rangle$, so it is finite-dimensional.

Example The space $\mathcal{M}_{2\times 2}$ of 2×2 matrices is finite-dimensional. Here is one basis with finitely many members.

$$\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rangle$$

Note From this point on we will restrict our attention to vector spaces that are finite-dimensional. All the later examples and definitions assume this of the spaces.

We will show that for any finite-dimensional space, all of its bases have the same number of elements.

Example Each of these is a basis for $\mathcal{P}_2$.

$$\begin{aligned} B_0 &= \langle 1, 1+x, 1+x+x^2 \rangle \\ B_1 &= \langle 1+x+x^2, 1+x, 1 \rangle \\ B_2 &= \langle x^2, 1+x, 1-x \rangle \\ B_3 &= \langle 1, x, x^2 \rangle \end{aligned}$$

Each has three elements.

Example Here are two different bases for $\mathcal{M}_{2\times 2}$.

$$B_0 = \langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rangle$$

$$B_1 = \langle \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \rangle$$

Both have four elements.

Vector Spaces and Linear Systems

Rowspace

Example The matrix before Gauss's Method and the matrix after have equal row spaces.

$$M = \begin{pmatrix} 1 & 2 & 1 & 0 & 3 \\ -1 & -2 & 2 & 2 & 0 \\ 2 & 4 & 5 & 2 & 9 \end{pmatrix} \xrightarrow[-2\rho_1+\rho_3]{\rho_1+\rho_2} \xrightarrow{-\rho_2+\rho_3} \begin{pmatrix} 1 & 2 & 1 & 0 & 3 \\ 0 & 0 & 3 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The nonzero rows of the latter matrix form a basis for Rowspace(M).

$$B = \langle (1 \ \ 2 \ \ 1 \ \ 0 \ \ 3), (0 \ \ 0 \ \ 3 \ \ 2 \ \ 3) \rangle$$

The row rank is 2.

So Gauss's Method produces a basis for the row space of a matrix. It has found the "repeat" information, that M's third row is three times the first plus the second, and eliminated that extra row.

Transpose I

?? *Definition* The *transpose* of a matrix is the result of interchanging its rows and columns, so that column j of the matrix A is row j of A^T and vice versa.

Example To find a basis for the column space of a matrix,

$$\begin{pmatrix} 2 & 3 \\ -1 & 1/2 \end{pmatrix}$$

transpose,

$$\begin{pmatrix} 2 & 3 \\ -1 & 1/2 \end{pmatrix}^T = \begin{pmatrix} 2 & -1 \\ 3 & 1/2 \end{pmatrix}$$

reduce,

$$\begin{pmatrix} 2 & -1 \\ 3 & 1/2 \end{pmatrix} \xrightarrow{(-3/2)\rho_1+\rho_2} \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}$$

and transpose back.

$$\begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}^T = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}$$

Transpose II

This basis

$$B = \langle \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \rangle$$

shows that the column space is the entire vector space $\mathbb{R}^2$.

Rank

?? *Definition* The *rank* of a matrix is its row rank or column rank.

Example The column rank of this matrix

$$\begin{pmatrix} 2 & -1 & 3 & 1 & 0 & 1 \\ 3 & 0 & 1 & 1 & 4 & -1 \\ 4 & -2 & 6 & 2 & 0 & 2 \\ 1 & 0 & 3 & 0 & 0 & 2 \end{pmatrix}$$

is 3. Its largest set of linearly independent columns is size 3 because that's the size of its largest set of linearly independent rows.

$$\xrightarrow[\substack{-2\rho_1+\rho_3 \\ -(1/2)\rho_1+\rho_4}]{-(3/2)\rho_1+\rho_2} \xrightarrow{-(1/3)\rho_2+\rho_4} \xrightarrow{\rho_3 \leftrightarrow \rho_4} \begin{pmatrix} 2 & -1 & 3 & 1 & 0 & 1 \\ 0 & 3/2 & -7/2 & -1/2 & 4 & -5/2 \\ 0 & 0 & 8/3 & -1/3 & -4/3 & 7/3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

?? *Theorem* For linear systems with n unknowns and with matrix of coefficients A, the statements

(1) the rank of A is r

(2) the vector space of solutions of the associated homogeneous system has dimension $n - r$

are equivalent.

Proof The rank of A is r if and only if Gaussian reduction on A ends with r nonzero rows. That's true if and only if echelon form matrices row equivalent to A have r-many leading variables. That in turn holds if and only if there are $n - r$ free variables. QED

Sums and Scalar Products

Definition of matrix sum and scalar multiple

?? *Definition* The *scalar multiple* of a matrix is the result of entry-by-entry scalar multiplication. The *sum* of two same-sized matrices is their entry-by-entry sum.

Example Where

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 2 \\ 9 & -1/2 & 5 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 \\ 8 & -1 \end{pmatrix}$$

Then

$$A + C = \begin{pmatrix} 2 & -1 \\ 10 & 2 \end{pmatrix} \qquad 5B = \begin{pmatrix} 0 & 0 & 10 \\ 45 & -5/2 & 25 \end{pmatrix}$$

Note that none of these is defined: $A + B$, $B + A$, $B + C$, $C + B$.

From the definition, they are not defined because the sizes don't match and so the entry-by-entry sum is not possible. But really they are not defined because the underlying function operations are not possible. The fact that A has two columns means that functions represented by A have two-dimensional domains. Functions represented by B have three-dimensional domains. Adding the two functions would be adding apples and oranges.

Matrix Multiplication

Example Consider two linear functions $h\colon V \to W$ and $g\colon W \to X$ represented as here.

$$\mathrm{Rep}_{B,C}(h) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \qquad \mathrm{Rep}_{C,D}(g) = \begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix}$$

We will do an explatory computation, to see how these two representations combine to give the representation of the composition $g \circ h\colon V \to X$.

Example Consider two linear functions $h: V \to W$ and $g: W \to X$ represented as here.

$$\mathrm{Rep}_{B,C}(h) = \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \qquad \mathrm{Rep}_{C,D}(g) = \begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix}$$

We will do an explatory computation, to see how these two representations combine to give the representation of the composition $g \circ h: V \to X$.

We start with the action of h on $\vec{v} \in V$.

$$\begin{aligned} \mathrm{Rep}_C(h(\vec{v})) &= \mathrm{Rep}_{B,C}(h) \cdot \mathrm{Rep}_B(\vec{v}) \\ &= \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 3v_1 + v_2 \\ 2v_1 + 5v_2 \\ 4v_1 + 6v_2 \end{pmatrix} \end{aligned}$$

Next, to that apply g.

$$\text{Rep}_{C,D}(g) \cdot \text{Rep}_C(h(\vec{v})) = \begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix} \begin{pmatrix} 3v_1 + v_2 \\ 2v_1 + 5v_2 \\ 4v_1 + 6v_2 \end{pmatrix}$$

$$= \begin{pmatrix} 8(3v_1 + v_2) + 7(2v_1 + 5v_2) + 11(4v_1 + 6v_2) \\ 9(3v_1 + v_2) + 10(2v_1 + 5v_2) + 12(4v_1 + 6v_2) \end{pmatrix}$$

Gather terms.

$$= \begin{pmatrix} (8 \cdot 3 + 7 \cdot 2 + 11 \cdot 4)v_1 + (8 \cdot 1 + 7 \cdot 5 + 11 \cdot 6)v_2 \\ (9 \cdot 3 + 10 \cdot 2 + 12 \cdot 4)v_1 + (9 \cdot 1 + 10 \cdot 5 + 12 \cdot 6)v_2 \end{pmatrix}$$

Rewrite as a matrix-vector multiplication.

$$= \begin{pmatrix} 8 \cdot 3 + 7 \cdot 2 + 11 \cdot 4 & 8 \cdot 1 + 7 \cdot 5 + 11 \cdot 6 \\ 9 \cdot 3 + 10 \cdot 2 + 12 \cdot 4 & 9 \cdot 1 + 10 \cdot 5 + 12 \cdot 6 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

So here is how the two starting matrices combine.

$$\begin{pmatrix} 8 & 7 & 11 \\ 9 & 10 & 12 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 5 \\ 4 & 6 \end{pmatrix} = \begin{pmatrix} 8 \cdot 3 + 7 \cdot 2 + 11 \cdot 4 & 8 \cdot 1 + 7 \cdot 5 + 11 \cdot 6 \\ 9 \cdot 3 + 10 \cdot 2 + 12 \cdot 4 & 9 \cdot 1 + 10 \cdot 5 + 12 \cdot 6 \end{pmatrix}$$

Definition of matrix multiplication

?? *Definition* The *matrix-multiplicative product* of the $m \times r$ matrix G and the $r \times n$ matrix H is the $m \times n$ matrix P, where

$$p_{i,j} = g_{i,1}h_{1,j} + g_{i,2}h_{2,j} + \cdots + g_{i,r}h_{r,j}$$

so that the i,j-th entry of the product is the dot product of the i-th row of the first matrix with the j-th column of the second.

$$GH = \begin{pmatrix} & \vdots & & \\ g_{i,1} & g_{i,2} & \cdots & g_{i,r} \\ & \vdots & & \end{pmatrix} \begin{pmatrix} & h_{1,j} & \\ \cdots & h_{2,j} & \cdots \\ & \vdots & \\ & h_{r,j} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & p_{i,j} & \cdots \\ & \vdots & \end{pmatrix}$$

Example

$$\begin{pmatrix} 3 & 1 & 6 \\ 2 & 5 & 9 \end{pmatrix} \begin{pmatrix} 2 & 0 & 4 \\ 1 & -3 & 5 \\ 4 & 2 & 7 \end{pmatrix} = \begin{pmatrix} 31 & 9 & 59 \\ 45 & 3 & 96 \end{pmatrix}$$

Example This product

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \end{pmatrix}$$

is not defined because the number of columns on the left must equal the number of rows on the right.

Example This product

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \end{pmatrix}$$

is not defined because the number of columns on the left must equal the number of rows on the right.

Example Square matrices of the same size have a defined product.

$$\begin{pmatrix} 1 & 3 & -1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 7 & 1 \\ 2 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 14 & -1 \\ 0 & 0 & 0 \\ 10 & 14 & 2 \end{pmatrix}$$

This reflects the fact that we can compose two functions from a space to itself $g, h \colon V \to V$.

Order, dimensions, and sizes

An important observation about the order in which we write these things: in writing the composition $g \circ h$, the function g is written first, that is, leftmost, but it is applied second.

$$\vec{v} \overset{h}{\longmapsto} h(\vec{v}) \overset{g}{\longmapsto} g(h(\vec{v}))$$

That order carries over to matrices: $g \circ h$ is represented by GH.

Also consider the dimensions of the spaces.

$$\text{dimension } n \text{ space} \xrightarrow{h} \text{dimension } r \text{ space} \xrightarrow{g} \text{dimension } m \text{ space}$$

Briefly, $m \times r$ times $r \times n$ equals $m \times n$, as here.

$$\begin{array}{cccc} 2\times 3 & 3\times 4 & = & 2\times 4 \\ \begin{pmatrix} 2 & 1 & 4 \\ -1 & 0 & 3 \end{pmatrix} & \begin{pmatrix} 3 & 0 & 2 & 1 \\ 5 & 0 & 0 & 2 \\ 1 & -1 & 4 & 7 \end{pmatrix} & = & \begin{pmatrix} 15 & -4 & 20 & 32 \\ 0 & -3 & 10 & 20 \end{pmatrix} \end{array}$$

Matrix multiplication is not commutative

Function composition is in general not a commutative operation — $\cos(\sqrt{x})$ is different than $\sqrt{\cos(x)}$. This holds even in the special case of composition of linear functions.

Matrix multiplication is not commutative

Function composition is in general not a commutative operation — $\cos(\sqrt{x})$ is different than $\sqrt{\cos(x)}$. This holds even in the special case of composition of linear functions.

Example Changing the order in which we multiply these matrices

$$\begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 12 & 33 \\ 24 & 20 \end{pmatrix}$$

changes the result.

$$\begin{pmatrix} -2 & 6 \\ 6 & 5 \end{pmatrix} \begin{pmatrix} 3 & 3 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} -6 & 18 \\ 18 & 38 \end{pmatrix}$$

Example The product of these two is defined in one order and not defined in the other.

$$\begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 8 & 12 & 0 \\ -4 & 0 & 1/2 \end{pmatrix}$$

Inverses

Definition of matrix inverse

?? *Definition* A matrix G is a *left inverse matrix* of the matrix H if GH is the identity matrix. It is a *right inverse* if HG is the identity. A matrix H with a two-sided inverse is an *invertible matrix*. That two-sided inverse is denoted H^{-1}.

Definition of matrix inverse

?? *Definition* A matrix G is a *left inverse matrix* of the matrix H if GH is the identity matrix. It is a *right inverse* if HG is the identity. A matrix H with a two-sided inverse is an *invertible matrix*. That two-sided inverse is denoted H^{-1}.

Example This matrix

$$H = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$$

has a two-sided inverse.

$$H^{-1} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

To check that, we multiply them in both orders. Here is one; the other is just as easy.

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example One advantage of knowing a matrix inverse is that it makes solving a linear system easy and quick. To solve

$$\begin{aligned} 2x + 5y &= -3 \\ x + 3y &= \ 10 \end{aligned}$$

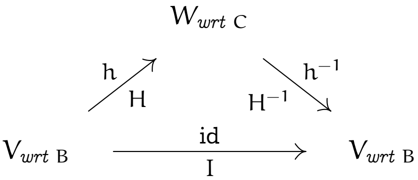
rewrite as a matrix equation

$$\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3 \\ 10 \end{pmatrix}$$

and multiply both sides (from the left) by the matrix inverse.

$$\begin{aligned} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 10 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -59 \\ 23 \end{pmatrix} \\ \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -59 \\ 23 \end{pmatrix} \end{aligned}$$

This specializes the arrow diagram for composition to the case of inverses.



?? *Lemma* If a matrix has both a left inverse and a right inverse then the two are equal.

?? *Theorem* A matrix is invertible if and only if it is nonsingular.

Proof *(For both results.)* Given a matrix H, fix spaces of appropriate dimension for the domain and codomain and fix bases for these spaces. With respect to these bases, H represents a map h. The statements are true about the map and therefore they are true about the matrix. QED

Finding the inverse of a matrix A is a lot of work but as we noted earlier, once we have it then solving linear systems $A\vec{x} = \vec{b}$ is easy.

Example The linear system

$$\begin{aligned} x + 3y + z &= 2 \\ 2x \qquad - z &= 12 \\ x + 2y \qquad &= 4 \end{aligned}$$

is this matrix equation.

$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 12 \\ 4 \end{pmatrix}$$

Solve it by multiplying both sides from the left by the inverse that we found earlier.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/3 & 2/3 & -1 \\ -1/3 & -1/3 & 1 \\ 4/3 & 1/3 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 12 \\ 4 \end{pmatrix} = \begin{pmatrix} 16/3 \\ -2/3 \\ -4/3 \end{pmatrix}$$

We sometimes want to repeatedly solve systems with the same left side but different right sides. This system equals the one on the prior slide but for one number on the right.

$$\begin{aligned} x + 3y + z &= 1 \\ 2x \qquad - z &= 12 \\ x + 2y \qquad &= 4 \end{aligned}$$

The solution is this.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2/3 & 2/3 & -1 \\ -1/3 & -1/3 & 1 \\ 4/3 & 1/3 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 12 \\ 4 \end{pmatrix} = \begin{pmatrix} 14/3 \\ -1/3 \\ -8/3 \end{pmatrix}$$

The inverse of a 2×2 matrix

?? *Corollary* The inverse for a 2×2 matrix exists and equals

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

if and only if $ad - bc \neq 0$.

Example

$$\begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix}^{-1} = \frac{1}{6} \begin{pmatrix} 1 & -4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/6 & -2/3 \\ 1/6 & 1/3 \end{pmatrix}$$

Properties of Determinants

Nonsingular matrices

For any matrix, whether or not it is nonsingular is a key question. Recall that an $n \times n$ matrix T is nonsingular if and only if each of these holds:

- any system $T\vec{x} = \vec{b}$ has a solution and that solution is unique;
- Gauss-Jordan reduction of T yields an identity matrix;
- the rows of T form a linearly independent set;
- the columns of T form a linearly independent set, a basis for $\mathbb{R}^n$;
- (any map that T represents is an isomorphism;)
- an inverse matrix T^{-1} exists.

This chapter develops a formula to determine whether a matrix is nonsingular.

Determining nonsingularity is trivial for 1×1 matrices.

$$\begin{pmatrix} a \end{pmatrix} \quad \text{is nonsingular iff} \quad a \neq 0$$

For the 2×2 formula.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{is nonsingular iff} \quad ad - bc \neq 0$$

Formula for the 3×3 case

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \text{is nonsingular iff} \quad aei + bfg + cdh - hfa - idb - gec \neq 0$$

With these cases in mind, we posit a family of formulas: a, $ad - bc$, etc. For each n the formula defines a *determinant* function $\det_{n\times n} : \mathcal{M}_{n\times n} \to \mathbb{R}$ such that an $n\times n$ matrix T is nonsingular if and only if $\det_{n\times n}(T) \neq 0$.

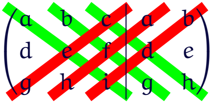
Warning

The formula for the determinant of a 2×2 matrix has something to do with multiplying diagonals.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Sometimes people have learned a mnemonic for the 3×3 formula that has to do with multplying diagonals.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb$$

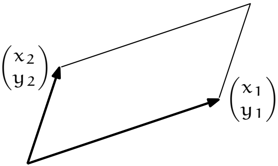


Don't try to extend to 4×4 or larger sizes; there is no general pattern here. Instead, for larger matrices use Gauss's Method.

Determinants as size functions

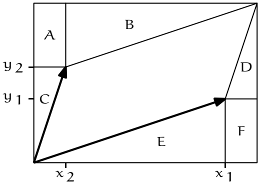
Box

This parallelogram is defined by the two vectors.



?? *Definition* In $\mathbb{R}^n$ the *box* (or *parallelepiped*) formed by $\langle \vec{v}_1, \ldots, \vec{v}_n \rangle$ is the set $\{t_1\vec{v}_1 + \cdots + t_n\vec{v}_n \mid t_1, \ldots, t_n \in [0 \ldots 1]\}$.

Area



$$
\begin{aligned}
\text{box area} &= \text{rectangle area} - \text{area of } A - \cdots - \text{area of } F \\
&= (x_1 + x_2)(y_1 + y_2) - x_2 y_1 - x_1 y_1/2 \\
&\quad - x_2 y_2/2 - x_2 y_2/2 - x_1 y_1/2 - x_2 y_1 \\
&= x_1 y_2 - x_2 y_1
\end{aligned}
$$

The determinant of this matrix gives the size of the box formed by the matrix's columns.

$$
\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1
$$

Determinants are multiplicative

?? *Theorem* A transformation $t\colon \mathbb{R}^n \to \mathbb{R}^n$ changes the size of all boxes by the same factor, namely, the size of the image of a box $|t(S)|$ is $|T|$ times the size of the box $|S|$, where T is the matrix representing t with respect to the standard basis.

That is, the determinant of a product is the product of the determinants $|TS| = |T| \cdot |S|$.

Example The transformation $t_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ that rotates all vectors through a counterclockwise angle θ is represented by this matrix.

$$T_\theta = \text{Rep}_{\mathcal{E}_2,\mathcal{E}_2}(t_\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Observe that t_θ doesn't change the size of any boxes, it just rotates the entire box as a rigid whole. Note that $|T_\theta| = 1$.

Determinant of the inverse

?? *Corollary* If a matrix is invertible then the determinant of its inverse is the inverse of its determinant $|T^{-1}| = 1/|T|$.

Proof $1 = |I| = |TT^{-1}| = |T| \cdot |T^{-1}|$ QED

Example These matrices are inverse.

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2 \qquad \begin{vmatrix} -2 & 1 \\ 3/2 & -1/2 \end{vmatrix} = -1/2$$

Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors

?? *Definition* A transformation $t\colon V \to V$ has a scalar *eigenvalue* λ if there is a nonzero *eigenvector* $\vec{\zeta} \in V$ such that $t(\vec{\zeta}) = \lambda \cdot \vec{\zeta}$.

Eigenvalues and eigenvectors

?? *Definition* A transformation $t\colon V \to V$ has a scalar *eigenvalue* λ if there is a nonzero *eigenvector* $\vec{\zeta} \in V$ such that $t(\vec{\zeta}) = \lambda \cdot \vec{\zeta}$.

?? *Definition* A square matrix T has a scalar *eigenvalue* λ associated with the nonzero *eigenvector* $\vec{\zeta}$ if $T\vec{\zeta} = \lambda \cdot \vec{\zeta}$.

Example The matrix

$$D = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

has an eigenvalue $\lambda_1 = 4$ and a second eigenvalue $\lambda_2 = 2$. The first is true because an associated eigenvector is $\vec{e_1}$

$$\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 4 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and similarly for the second an associated eigenvector is e_2.

$$\begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Computing eigenvalues and eigenvectors

Example We will find the eigenvalues and associated eigenvectors of this matrix.

$$T = \begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix}$$

We want to find scalars x such that $T\vec{\zeta} = x\vec{\zeta}$ for some nonzero $\vec{\zeta}$. Bring the terms to the left side.

$$\begin{pmatrix} 0 & 5 & 7 \\ -2 & 7 & 7 \\ -1 & 1 & 4 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} - x \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and factor.

$$\begin{pmatrix} 0-x & 5 & 7 \\ -2 & 7-x & 7 \\ -1 & 1 & 4-x \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (*)$$

This homogeneous system has nonzero solutions if and only if the matrix is singular, that is, has a determinant of zero.

Some computation gives the determinant and its factors.

$$\begin{aligned}0 &= \begin{vmatrix} 0-x & 5 & 7 \\ -2 & 7-x & 7 \\ -1 & 1 & 4-x \end{vmatrix} \\ &= x^3 - 11x^2 + 38x - 40 = (x-5)(x-4)(x-2)\end{aligned}$$

So the eigenvalues are $\lambda_1 = 5$, $\lambda_2 = 4$, and $\lambda_3 = 2$.

Characteristic polynomial

?? *Definition* The *characteristic polynomial of a square matrix* T is the determinant $|T - xI|$ where x is a variable. The *characteristic equation* is $|T - xI| = 0$. The *characteristic polynomial of a transformation* t is the characteristic polynomial of any matrix representation $\mathrm{Rep}_{B,B}(t)$.

A criteria for diagonalizability

?? *Corollary* An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Matrix Exponentials

Matrix Exponentials I

Let M be an $n \times n$ real or complex matrix. The exponential of M, denoted by e^M or $\exp(M)$, is the $n \times n$ matrix given by the power series

$$e^M = \sum_{k=0}^{\infty} \frac{1}{k!} M^k \tag{1}$$

where M^0 is defined to be the identity matrix I with the same dimensions as M.

Properties

- $e^Z = I$
- $\exp(M^T) = \exp(M)^T$, where M^T denotes the transpose of M
- $exp(M^\star) = exp(M)^\star$, where $M^\star$ denotes the conjugate transpose of M
- If K is invertible then $\exp(KMK^{-1}) = K\exp(MK^{-1})$
- If $MK = KM$ then $e^M e^K = e^{M+K}$

Matrix Exponentials II

The proof of this identity is the same as the standard power-series argument for the corresponding identity for the exponential of real numbers. That is to say, as long as M and K commute, it makes no difference to the argument whether M and K are numbers or matrices.

Consequences of the preceding identity are the following:

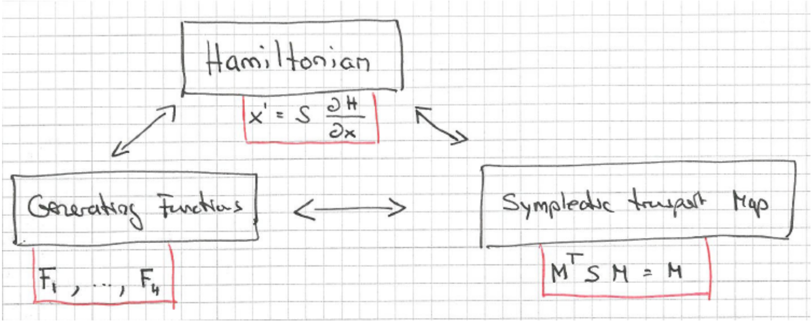
- $e^{aM}e^{bM} = e^{(a+b)M}$ for $a, b \in \mathbb{R}$
- $e^{M}e^{-M} = I$

Here a few important relations to remember: if M is symmetric then e^M is also symmetric, and if M is skew-symmetric then e^M is orthogonal. If M is Hermitian then e^M is also Hermitian, and if M is skew-Hermitian then e^M is unitary.

Remember: a Hermitian matrix (or self-adjoint matrix) is a complex square matrix that is equal to its own conjugate transpose.

The Symplectic Form of Hamilton's EQM

The Symplectic Form of Hamilton's EQM I



The terms **canonical transformation** and **symplectic condition** refer to transformations of the form

$$J^T S J = S. \tag{2}$$

and explicitly expresses the **symplectic** condition. It is a constraint upon the Jacobian matrix J of a canonical transformation. In fact, it is a necessary and sufficient condition for all canonical transformations.

As a consequence:

The solutions of Hamilton's equations of motion or equivalently the motions of Hamiltonian systems are themselves canonical transformations and therefore obey the symplectic condition.

Symplectic Matrices and Groups I

In a purely mathematical sense, (??) defines a set of even-dimensioned square matrices which form a group (called appropriately the symplectic group $Sp(2N)$ just as the associated canonical transformations form a group.

Since for every canonical transformation, there exists a corresponding symplectic (Jacobian) matrix, the group property of the canonical transformations implies the group property of the matrices. However it is straightforward to demonstrate the group properties of the matrices directly by verifying that they satisfy the following conditions

1) $(J_1 J_2)J_3 = J_1(J_2 J_3)$ (associative law)
2) Every product of two elements and the square of each element is a member of the group.
3) The group contains the (symplectic) unit element.
4) $J^{-1}J = 1$. Since their determinants do not vanish, all symplectic matrices have inverses.

Symplectic Matrices and Groups II

The symplectic condition (??) constitutes a fairly substantial set of constraints, one of the most important of which concerns the determinant of the matrix. Since the determinant of the product of two square matrices of equal dimension is equal to the product of their determinants, the square of the determinant of a symplectic matrix, and therefore of the Jacobian matrix of a canonical transformation, is equal to

$$\det \mathrm{J}^2 = |\mathrm{J}|^2 = 1. \tag{3}$$

Equation (??) is a main factor in demonstrating Liouville's Theorem.

For a two-by-two symplectic matrix, the determinant condition is the sole constraint; it completely determines the symplecticity of the matrix.

For a single degree of freedom, uncoupled to any other, the determinant condition is equivalent to the symplectic condition.

Symplectic Matrices and Groups III

For matrices of higher order than two, the symplectic condition implies more constraints among the matrix elements than the one expressed by (??). It turns out that the number of independent matrix elements in a symplectic matrix of order $2N$ is $N(2N+1)$, so the number of constraints among the $(2N)^2$ elements is

$$\text{number of constraints} : N(2N-1).$$

It can be shown (exercise) that any real symplectic matrix J sufficiently near the identity can be written

$$J = \exp SA_s, \tag{4}$$

where A_s is a real symmetric matrix and S is of the same order as A_s. For matrices near the identity, this establishes a **injective** (one-to-one) relationship between real symplectic matrices and real symmetric matrices.

Eigenvalues of J I

In beam dynamics, we are primarily interested in real coordinate systems, and the Jacobian matrices of transformations between real systems are real matrices. Consequently the symplectic matrices we deal with are also real, and the eigenvalues of real, symplectic matrices form quite a restricted set [Dragt]. The eigenvalues are the roots of the characteristic polynomial,

$$P(\lambda) = |J - \lambda 1|. \tag{5}$$

We note that $P(\lambda)$ has degree $2N$ and the characteristic equation is

$$P(\lambda) = 0. \tag{6}$$

We immediately can state that $\lambda = 0$ cannot be a root and remember that this equation has $2N$ roots, and the product of the roots is equal to the constant term of the polynomial.

$$\prod_{i=1}^{2N} \lambda_i = \det J = 1. \tag{7}$$

Eigenvalues of J II

Either the eigenvalues are real or they occur in complex conjugate pairs. Appealing only to the properties of symplectic matrices, it can also be shown that

$$\lambda^{-N} P(\lambda) = \lambda^{N} P\left(\frac{1}{\lambda}\right), \tag{8}$$

with the consequence that if λ is a root, so is its reciprocal $1/\lambda$, and furthermore, these roots have the same multiplicities ($+1$ or -1), then that root has even multiplicity. Properties of the λ:

1) They are real or they occur in complex conjugate pairs.
2) They occur in reciprocal pairs, each member of a pair having the same multiplicity.
3) If either $+1$ or -1 is an eigenvalue, it has even multiplicity.

When combined, the conditions just enumerated place strong restrictions on the possible eigenvalues of a real symplectic matrix.

- the the eigenvalues cannot all lie inside or all lie outside the unit circle
- from a dynamics system point of view, we can say that the linear part of a symplectic map at a fixed point is a symplectic matrix

Eigenvalues of the 2×2 Case I

Consider first the simplest case of a 2×2 real symplectic matrix with eigenvalues λ_1 and λ_2. With the reciprocal property, we get immediately

$$\lambda_1 \lambda_2 = 1 \tag{9}$$

- $\lambda_1, \lambda_2 \in \mathbb{R}$
  - Suppose, now, that $\lambda_1 > 1 :\to \quad \lambda_2 < 1$.
  - Similarly, if $\lambda_1 < -1 :\to \qquad 0 > \lambda_2 > -1$
- $\lambda_1 \in \mathbb{C}$
  - $\lambda_2 = \bar{\lambda}_1$

This condition, when combined with (??), shows that in this case λ_1 and λ_2 must lie on the unit circle in the complex plane. Finally, there are the two special cases $\lambda_1 = \lambda_2 = 1$ and $\lambda_1 = \lambda_2 = -1$ Altogether, there are five possible cases. They are listed below along with names and designations whose significance will become clear later on.

Eigenvalues of the 2×2 Case II

1) Hyperbolic case (unstable): $\lambda_1 > 1$ and $0 < \lambda_2 < 1$.
2) Inversion hyperbolic case (unstable): $\lambda_1 < -1$ and $-1 < \lambda_2 < 0$.
3) Elliptic case (stable): $\lambda_1 = \exp i\phi, \lambda_1 = \exp -i\phi$. (Eigenvalues are complex conj. and lie on the unit circle).
4) Parabolic case (generally linearly unstable): $\lambda_1 = \lambda_2 = 1$.
5) Inversion parabolic case (generally linearly unstable): $\lambda_1 = \lambda_2 = -1$.

Note that in all cases both eigenvalues cannot lie inside the unit circle nor can both eigenvalues lie outside the unit circle.

Eigenvalues of the 2×2 Case III

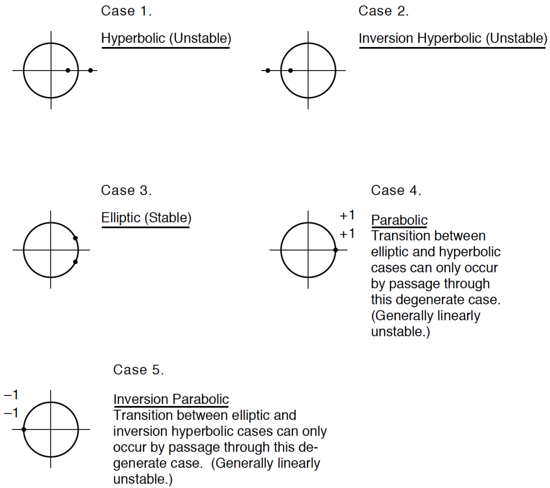


Figure : Possible cases for the eigenvalues of a $2 \times 2, \mathbb{R}$ symplectic matrix [Dragt].

Symplectic Condition and Poisson Brackets I

What is remarkable is that the symplectic condition

$$J^T S J = S.$$

applies also to a nonlinear system if we identify J to be the Jacobian matrix of the map, whose elements are defined as

$$J_{i,j} = \frac{\partial X_i}{\partial (X_0)_j} \tag{10}$$

where $(X_0)_j$ is the j-th component of the initial coordinates of a particle at $s = 0$, X_i is the i-th component of the final state X of the particle at an arbitrary position s. In a linear system, the Jacobian matrix is just the transformation matrix, and is independent of the particle coordinates.

In a nonlinear system, the Jacobian matrix J depends on the components of X_0, and the symplectic condition must be satisfied for all X_0.

Symplectic Condition and Poisson Brackets II

- Another consequence of a symplectic map is that it obeys the **Liouville theorem**, i.e. the phase space volume is conserved as the system evolves according to the map.
- Symplectic maps therefore are area-preserving maps.
- Liouville theorem follows because the Jacobian matrix, being symplectic, has unit determinant, which in turn assures that a volume element in phase space maintains its volume as it evolves with time.

If time permits: Similarity Definition and Examples

Similar matrices

?? *Definition* The matrices T and $\hat{T}$ are *similar* if there is a nonsingular P such that $\hat{T} = PTP^{-1}$.

Example Consider the derivative map $d/dx\colon \mathcal{P}_2 \to \mathcal{P}_2$. Fix the basis $B = \langle 1, x, x^2 \rangle$ and the basis $D = \langle 1, 1+x, 1+x+x^2 \rangle$. In this arrow diagram we will first get T, and then calculate $\hat{T}$ from it.

$$\begin{array}{ccc} V_{\mathit{wrt}\ B} & \xrightarrow[T]{t} & V_{\mathit{wrt}\ B} \\ {\scriptstyle \mathrm{id}}\big\downarrow & & {\scriptstyle \mathrm{id}}\big\downarrow \\ V_{\mathit{wrt}\ D} & \xrightarrow[\hat{T}]{t} & V_{\mathit{wrt}\ D} \end{array}$$

Similar matrices

?? *Definition* The matrices T and $\hat{T}$ are *similar* if there is a nonsingular P such that $\hat{T} = PTP^{-1}$.

Example Consider the derivative map $d/dx\colon \mathcal{P}_2 \to \mathcal{P}_2$. Fix the basis $B = \langle 1, x, x^2 \rangle$ and the basis $D = \langle 1, 1+x, 1+x+x^2 \rangle$. In this arrow diagram we will first get T, and then calculate $\hat{T}$ from it.

$$\begin{array}{ccc} V_{\mathit{wrt}\ B} & \xrightarrow[T]{t} & V_{\mathit{wrt}\ B} \\ \scriptstyle{\text{id}}\big\downarrow & & \scriptstyle{\text{id}}\big\downarrow \\ V_{\mathit{wrt}\ D} & \xrightarrow[\hat{T}]{t} & V_{\mathit{wrt}\ D} \end{array}$$

The action of d/dx on the elements of the basis B is $1 \mapsto 0$, $x \mapsto 1$, and $x^2 \mapsto 2x$.

$$\text{Rep}_B(\frac{d}{dx}(1)) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{Rep}_B(\frac{d}{dx}(x)) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{Rep}_B(\frac{d}{dx}(x^2)) = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

So we have this matrix representation of the map.

$$T = \mathrm{Rep}_{B,B}(\frac{d}{dx}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

The matrix changing bases from B to D is $\mathrm{Rep}_{B,D}(\mathrm{id})$. We find these by eye

$$\mathrm{Rep}_D(\mathrm{id}(1)) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(\mathrm{id}(x)) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad \mathrm{Rep}_D(\mathrm{id}(x^2)) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

to get this.

$$P = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Now, by following the arrow diagram we have $\hat{T} = PTP^{-1}$.

$$\hat{T} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

To check that, and to underline what the arrow diagram says

$$\begin{array}{ccc} V_{wrt\ B} & \xrightarrow[T]{t} & V_{wrt\ B} \\ {\scriptstyle id}\big\downarrow & & {\scriptstyle id}\big\downarrow \\ V_{wrt\ D} & \xrightarrow[\hat{T}]{t} & V_{wrt\ D} \end{array}$$

we calculate $\hat{T}$ directly. The effect of the map on the basis elements is $d/dx(1) = 0$, $d/dx(1+x) = 1$, and $d/dx(1+x+x^2) = 1+2x$. Representing of those with respect to D

$$\text{Rep}_D(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{Rep}_D(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{Rep}_D(1+2x) = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

gives the same matrix $\hat{T} = \text{Rep}_{D,D}(d/dx)$ as above.

The definition doesn't require that we consider the underlying maps. We can just multiply matrices.

Example Where

$$T = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 4 & 5 & 2 \end{pmatrix} \qquad P = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

(note that P is nonsingular) we can compute this $\hat{T} = PTP^{-1}$.

$$\hat{T} = \begin{pmatrix} 2 & 0 & 0 \\ 3 & 1 & 4/3 \\ 27/2 & 3/2 & 2 \end{pmatrix}$$