Title

D. Zack Garza

Contents

1 Linear Algebra

Remark 1.0.1: The underlying field will be assumed to be $\mathbb{R}$ for this section.

1.1 Notation

$\mathrm{Mat}(m,n)$	the space of all $m \times n$ matrices
T	a linear map $\mathbb{R}^n \to \mathbb{R}^m$
$A \in \mathrm{Mat}(m,n)$	an $m \times n$ matrix representing T
$A^t \in \mathrm{Mat}(n,m)$	an $n \times m$ transposed matrix
$\mathbf{a}$	a $1 \times n$ column vector
$\mathbf{a}^t$	an $n \times 1$ row vector
$A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$	a matrix formed with $\mathbf{a}_i$ as the columns
V, W	vector spaces
$\|V\|, \dim(W)$	dimensions of vector spaces
$\det(A)$	the determinant of A
$\left[A \mid \mathbf{b}\right] \coloneqq [\mathbf{a}_1, \mathbf{a}_2, \cdots \mathbf{a}_n, \mathbf{b}]$	augmented matrices
$\left[A \mid B\right] \coloneqq [\mathbf{a}_1, \cdots \mathbf{a}_n, \mathbf{b}_1, \cdots, \mathbf{b}_m]$	block matrices
$\mathrm{Spec}(A)$	the multiset of eigenvalues of A
$A\mathbf{x} = \mathbf{b}$	a system of linear equations
$r \coloneqq \mathrm{rank}(A)$	the rank of A
$r_b = \mathrm{rank}\left(\left[A \mid \mathbf{b}\right]\right)$	the rank of A augmented by $\mathbf{b}$.

1.2 Big Theorems

Theorem 1.2.1 ***(Rank-Nullity).***

$$|\ker(A)| + |\mathrm{im}(A)| = |\mathrm{dom}(A)|,$$

where $\mathrm{nullspace}(A) = |\mathrm{im}\, A|, \mathrm{rank}(A) = |\mathrm{im}(A)|$, and n is the number of columns in the corresponding matrix.
Generalization: the following sequence is always exact:

$$0 \to \ker(A) \overset{\mathrm{id}}{\hookrightarrow} \mathrm{dom}(A) \overset{A}{\twoheadrightarrow} \mathrm{im}(A) \to 0.$$

Moreover, it always splits, so $\mathrm{dom}\, A = \ker A \oplus \mathrm{im}\, A$ and thus $|\mathrm{dom}(A)| = |\ker(A)| + |\mathrm{im}(A)|$.

Remark 1.2.1: We also have

$$\dim(\mathrm{rowspace}(A)) = \dim(\mathrm{colspace}(A)) = \mathrm{rank}(A).$$

1.3 Big List of Equivalent Properties

Let A be an $m \times n$ matrix. TFAE: - A is invertible and has a unique inverse A^{-1} - A^T is invertible - $\det(A) \neq 0$ - The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $b \in \mathbb{R}^m$ - The homogeneous system $A\mathbf{x} = 0$ has only the trivial solution $\mathbf{x} = 0$ - $\mathrm{rank}(A) = n$ - i.e. A is full rank - $\mathrm{nullity}(A) := \dim \mathrm{nullspace}(A) = 0$ - $A = \prod_{i=1}^{k} E_i$ for some finite k, where each E_i is an elementary matrix. - A is row-equivalent to the identity matrix I_n - A has exactly n pivots - The columns of A are a basis for $\mathbb{R}^n$ - i.e. $\mathrm{colspace}(A) = \mathbb{R}^n$ - The rows of A are a basis for $\mathbb{R}^m$ - i.e. $\mathrm{rowspace}(A) = \mathbb{R}^m$ - $(\mathrm{colspace}(A))^\perp = (\mathrm{rowspace} A)^\perp = \{\mathbf{0}\}$ - Zero is not an eigenvalue of A. - A has n linearly independent eigenvectors - The rows of A are coplanar.

Similarly, by taking negations, TFAE:

- A is not invertible
- A is singular
- A^T is not invertible
- $\det A = 0$
- The linear system $A\mathbf{x} = \mathbf{b}$ has either no solution or infinitely many solutions.
- The homogeneous system $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions
- $\mathrm{rank}\, A < n$
- $\dim \mathrm{nullspace}\ A > 0$
- At least one row of A is a linear combination of the others
- The *RREF* of A has a row of all zeros.

Reformulated in terms of linear maps T, TFAE: - $T^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ exists - $\mathrm{im}(T) = \mathbb{R}^n$ - $\ker(T) = 0$ - T is injective - T is surjective - T is an isomorphism - The system $A\mathbf{x} = 0$ has infinitely many solutions

1.4 Vector Spaces

1.4.1 Linear Transformations

Definition 1.4.1 (Linear Transformation)

todo

Remark 1.4.1: It is common to want to know the range and kernel of a specific linear transformation T. T can be given in many ways, but a general strategy for deducing these properties involves:

- Express an arbitrary vector in V as a linear combination of its basis vectors, and set it equal to an arbitrary vector in W.
- Use the linear properties of T to make a substitution from known transformations
- Find a restriction or relation given by the constants of the initial linear combination.

Remark 1.4.2: Useful fact: if $V \leq W$ is a subspace and $\dim(V) \geq \dim(W)$, then $V = W$.

Definition 1.4.2 (Kernel)

todo

Proposition 1.4.1 ***(Two-step vector subspace test).***
If $V \subseteq W$, then V is a subspace of W if the following hold:

$$\mathbf{0} \in V \tag{1}$$
$$\mathbf{a}, \mathbf{b} \in V \implies t\mathbf{a} + \mathbf{b} \in V. \tag{2}$$

1.4.2 Linear Independence

Proposition 1.4.2 ***(?).***
Any set of two vectors $\{\mathbf{v}, \mathbf{w}\}$ is linearly **dependent** $\iff \exists\lambda : \mathbf{v} = \lambda\mathbf{w}$, i.e. one is not a scalar multiple of the other.

1.4.3 Bases

Definition 1.4.3 (Basis and dimension)
A set S forms a **basis** for a vector space V iff

1. S is a set of linearly independent vectors, so $\sum \alpha_i \vec{s_i} = 0 \implies \alpha_i = 0$ for all i.
2. S spans V, so $\vec{v} \in V$ implies there exist α_i such that $\sum \alpha_i \vec{s_i} = \vec{v}$

In this case, we define the **dimension** of V to be $|S|$.

Show how to compute basis of kernel.

Show how to compute basis of row space (nonzero rows in **??**A)?)

Show how to compute basis of column space: leading ones.

1.4.4 The Inner Product

The point of this section is to show how an inner product can induce a notion of "angle", which agrees with our intuition in Euclidean spaces such as $\mathbb{R}^n$, but can be extended to much less intuitive things, like spaces of functions.

Definition 1.4.4 (The standard inner product)
The **Euclidean inner product** is defined as

$$\langle \mathbf{a},\ \mathbf{b} \rangle = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Also sometimes written as $\mathbf{a}^T \mathbf{b}$ or $\mathbf{a} \cdot \mathbf{b}$.

Proposition 1.4.3 ***(Inner products induce norms and angles).***
Yields a norm

$$\|\mathbf{x}\| \coloneqq \sqrt{\langle \mathbf{x},\ \mathbf{x} \rangle}$$

which has a useful alternative formulation

$$\langle \mathbf{x},\ \mathbf{x} \rangle = \|\mathbf{x}\|^2.$$

This leads to a notion of angle:

$$\langle \mathbf{x},\ \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta_{x,y} \implies \cos \theta_{x,y} \coloneqq \frac{\langle \mathbf{x},\ \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|} = \langle \widehat{\mathbf{x}},\ \widehat{\mathbf{y}} \rangle$$

where $\theta_{x,y}$ denotes the angle between the vectors $\mathbf{x}$ and $\mathbf{y}$.

Remark 1.4.3: Since $\cos\theta = 0$ exactly when $\theta = \pm\frac{\pi}{2}$, we can can declare two vectors to be **orthogonal** exactly in this case:

$$\mathbf{x} \in \mathbf{y}^{\perp} \iff \langle \mathbf{x},\ \mathbf{y} \rangle = 0.$$

Note that this makes the zero vector orthogonal to everything.

Definition 1.4.5 (Orthogonal Complement/Perp)
Given a subspace $S \subseteq V$, we define its **orthogonal complement**

$$S^{\perp} = \left\{ \mathbf{v} \in V \;\middle|\; \forall \mathbf{s} \in S,\ \langle \mathbf{v},\ \mathbf{s} \rangle = 0 \right\}.$$

Remark 1.4.4: Any choice of subspace $S \subseteq V$ yields a decomposition $V = S \oplus S^{\perp}$.

Proposition 1.4.4 ***(Formula expanding a norm and 'Pythagorean theorem').***
A useful formula is

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\langle \mathbf{x},\ \mathbf{y} \rangle + \|\mathbf{y}\|^2, .$$

When $\mathbf{x} \in \mathbf{y}^{\perp}$, this reduces to

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

Proposition 1.4.5 ***(Properties of the inner product).***

1. **Bilinearity:**

$$\left\langle \sum_j \alpha_j \mathbf{a}_j,\ \sum_k \beta_k \mathbf{b}_k \right\rangle = \sum_j \sum_i \alpha_j \beta_i \langle \mathbf{a}_j,\ \mathbf{b}_i \rangle.$$

2. **Symmetry**:

$$\langle \mathbf{a},\ \mathbf{b} \rangle = \langle \mathbf{b},\ \mathbf{a} \rangle$$

3. **Positivity**:

$$\mathbf{a} \neq \mathbf{0} \implies \langle \mathbf{a},\ \mathbf{a} \rangle > 0$$

4. **Nondegeneracy**:

$$\mathbf{a} = \mathbf{0} \iff \langle \mathbf{a},\ \mathbf{a} \rangle = 0$$

Proof of Cauchy-Schwarz: See Goode page 346.

1.4.5 Gram-Schmidt Process

Extending a basis $\{\mathbf{x}_i\}$ to an orthonormal basis $\{\mathbf{u}_i\}$

$$
\begin{aligned}
\mathbf{u}_1 &= N(\mathbf{x_1}) \\
\mathbf{u}_2 &= N(\mathbf{x}_2 - \langle \mathbf{x}_2,\ \mathbf{u}_1\rangle \mathbf{u}_1) \\
\mathbf{u}_3 &= N(\mathbf{x}_3 - \langle \mathbf{x}_3,\ \mathbf{u}_1\rangle \mathbf{u}_1 - \langle \mathbf{x}_3,\ \mathbf{u}_2\rangle \mathbf{u}_2) \\
\vdots \quad &\quad \vdots \\
\mathbf{u}_k &= N(\mathbf{x}_k - \sum_{i=1}^{k-1} \langle \mathbf{x}_k,\ \mathbf{u}_i\rangle \mathbf{u}_i)
\end{aligned}
$$

where N denotes normalizing the result.

In more detail The general setup here is that we are given an orthogonal basis $\{\mathbf{x}_i\}_{i=1}^n$ and we want to produce an **orthonormal** basis from them.

Why would we want such a thing? Recall that we often wanted to change from the standard basis $\mathcal{E}$ to some different basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots\}$. We could form the change of basis matrix $B = [\mathbf{b}_1, \mathbf{b}_2, \cdots]$ acts on vectors in the $\mathcal{B}$ basis according to

$$B[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{E}}.$$

But to change from $\mathcal{E}$ to $\mathcal{B}$ requires computing B^{-1}, which acts on vectors in the standard basis according to

$$B^{-1}[\mathbf{x}]_{\mathcal{E}} = [\mathbf{x}]_{\mathcal{B}}.$$

If, on the other hand, the $\mathbf{b}_i$ are orthonormal, then $B^{-1} = B^T$, which is much easier to compute. We also obtain a rather simple formula for the coordinates of $\mathbf{x}$ with respect to $\mathcal{B}$. This follows because we can write

$$\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{x},\ \mathbf{b}_i\rangle \mathbf{b}_i \coloneqq \sum_{i=1}^{n} c_i \mathbf{b}_i, .$$

and we find that

$$[\mathbf{x}]_{\mathcal{B}} = \mathbf{c} \coloneqq [c_1, c_2, \cdots, c_n]^T..$$

This also allows us to simplify projection matrices. Supposing that A has orthonormal columns and letting S be the column space of A, recall that the projection onto S is defined by

$$P_S = Q(Q^T Q)^{-1} Q^T.$$

Since Q has orthogonal columns and satisfies $Q^T Q = I$, this simplifies to

$$P_S = QQ^T..$$

The Algorithm Given the orthogonal basis $\{\mathbf{x}_i\}$, we form an orthonormal basis $\{\mathbf{u}_i\}$ iteratively as follows.

First define

$$\begin{aligned} N : \mathbb{R}^n &\to S^{n-1} \\ \mathbf{x} &\mapsto \widehat{\mathbf{x}} := \frac{\mathbf{x}}{\|\mathbf{x}\|} \end{aligned}$$

which projects a vector onto the unit sphere in $\mathbb{R}^n$ by normalizing. Then,

$$\begin{aligned} \mathbf{u}_1 &= N(\mathbf{x_1}) \\ \mathbf{u}_2 &= N(\mathbf{x}_2 - \langle \mathbf{x}_2,\ \mathbf{u}_1 \rangle \mathbf{u}_1) \\ \mathbf{u}_3 &= N(\mathbf{x}_3 - \langle \mathbf{x}_3,\ \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{x}_3,\ \mathbf{u}_2 \rangle \mathbf{u}_2) \\ \vdots \quad & \quad \vdots \\ \mathbf{u}_k &= N(\mathbf{x}_k - \sum_{i=1}^{k-1} \langle \mathbf{x}_k,\ \mathbf{u}_i \rangle \mathbf{u}_i) \end{aligned}$$

In words, at each stage, we take one of the original vectors $\mathbf{x}_i$, then subtract off its projections onto all of the $\mathbf{u}_i$ we've created up until that point. This leaves us with only the component of $\mathbf{x}_i$ that is orthogonal to the span of the previous $\mathbf{u}_i$ we already have, and we then normalize each $\mathbf{u}_i$ we obtain this way.

Alternative Explanation:

Given a basis

$$S = \{\mathbf{v_1}, \mathbf{v_2}, \cdots \mathbf{v_n}\},$$

the Gram-Schmidt process produces a corresponding orthogonal basis

$$S' = \{\mathbf{u_1}, \mathbf{u_2}, \cdots \mathbf{u_n}\}$$

that spans the same vector space as S.

S' is found using the following pattern:

$$\begin{aligned} \mathbf{u_1} &= \mathbf{v_1} \\ \mathbf{u_2} &= \mathbf{v_2} - \mathrm{proj}_{\mathbf{u_1}} \mathbf{v_2} \\ \mathbf{u_3} &= \mathbf{v_3} - \mathrm{proj}_{\mathbf{u_1}} \mathbf{v_3} - \mathrm{proj}_{\mathbf{u_2}} \mathbf{v_3} \end{aligned}$$

where

$$\text{proj}_{\mathbf{u}}\mathbf{v} = (\text{scal}_{\mathbf{u}}\mathbf{v})\frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{u}\|}\frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\langle \mathbf{v},\ \mathbf{u}\rangle}{\|\mathbf{u}\|^2}\mathbf{u}$$

is a vector defined as the *orthogonal projection of* $\mathbf{v}$ *onto* $\mathbf{u}$.

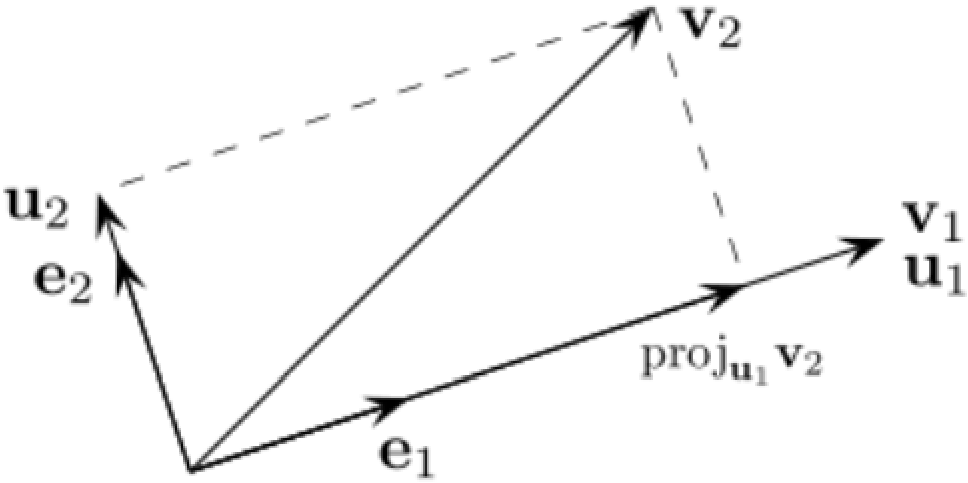


Figure 1: Image

The orthogonal set S' can then be transformed into an orthonormal set S'' by simply dividing the vectors $s \in S'$ by their magnitudes. The usual definition of a vector's magnitude is

$$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a},\ \mathbf{a}\rangle} \text{ and } \|\mathbf{a}\|^2 = \langle \mathbf{a},\ \mathbf{a}\rangle$$

As a final check, all vectors in S' should be orthogonal to each other, such that

$$\langle \mathbf{v_i},\ \mathbf{v_j}\rangle = 0 \text{ when } i \neq j$$

and all vectors in S'' should be orthonormal, such that

$$\langle \mathbf{v_i},\ \mathbf{v_j}\rangle = \delta_{ij}$$

1.4.6 The Fundamental Subspaces Theorem

Given a matrix $A \in \text{Mat}(m, n)$, and noting that

$$A : \mathbb{R}^n \to \mathbb{R}^m,$$
$$A^T : \mathbb{R}^m \to \mathbb{R}^n$$

We have the following decompositions:

$$\begin{aligned}
\mathbb{R}^n &\cong \ker A \oplus \operatorname{im} A^T &&\cong \operatorname{nullspace}(A) \oplus \operatorname{colspace}(A^T) \\
\mathbb{R}^m &\cong \operatorname{im} A \oplus \ker A^T &&\cong \operatorname{colspace}(A) \oplus \operatorname{nullspace}(A^T)
\end{aligned}$$

1.4.7 Computing change of basis matrices

todo

1.5 Matrices

Remark 1.5.1: An $m \times n$ matrix is a map from n-dimensional space to m-dimensional space. The number of *rows* tells you the dimension of the codomain, the number of *columns* tells you the dimension of the *domain*.

⚠**Warning 1.5.1:** The space of matrices is not an integral domain! Counterexample: if A is singular and nonzero, there is some nonzero $\mathbf{v}$ such that $A\mathbf{v} = \mathbf{0}$. Then setting $B = [\mathbf{v}, \mathbf{v}, \cdots]$ yields $AB = 0$ with $A \neq 0, B \neq 0$.

Definition 1.5.1 (Rank of a matrix)
The **rank** of a matrix A representing a linear transformation T is $\dim \operatorname{colspace}(A)$, or equivalently $\dim \operatorname{im} T$.

Proposition 1.5.1 ***(?).***
$\operatorname{rank}(A)$ is equal to the number of nonzero rows in $\operatorname{RREF}(A)$.

Definition 1.5.2 (Trace of a Matrix)

$$\operatorname{Trace}(A) = \sum_{i=1}^{m} A_{ii}$$

Definition 1.5.3 (Elementary Row Operations)
The following are **elementary row operations** on a matrix:

- Permute rows
- Multiple a row by a scalar
- Add any row to another

Proposition 1.5.2 ***(Formula for matrix multiplication).***
If $A = [\mathbf{a}_1, \mathbf{a}_2, \cdots] \in \text{Mat}(m, n)$ and $B = [\mathbf{b}_1, \mathbf{b}_2, \cdots] \in \text{Mat}(n, p)$, then

$$C := AB \implies c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = \langle \mathbf{a_i},\ \mathbf{b_j} \rangle$$

where $1 \leq i \leq m$ and $1 \leq j \leq p$. In words, each entry c_{ij} is obtained by dotting *row* i of A against *column* j of B.

1.5.1 Systems of Linear Equations

Definition 1.5.4 (Consistent and inconsistent)
A system of linear equations is **consistent** when it has at least one solution. The system is **inconsistent** when it has no solutions.

Definition 1.5.5 (Homogeneous Systems)
?

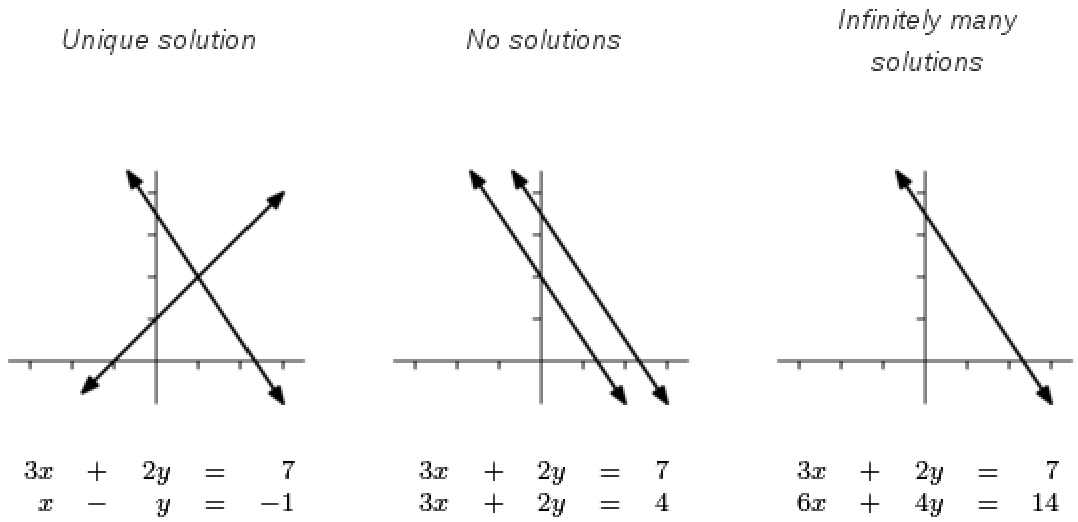
Remark 1.5.2: Homogeneous systems are always consistent, i.e. there is always at least one solution.

Remark 1.5.3:

- Tall matrices: more equations than unknowns, *overdetermined*
- Wide matrices: more unknowns than equations, *underdetermined*

Proposition 1.5.3 ***(Characterizing solutions to a system of linear equations).***
There are three possibilities for a system of linear equations:

1. No solutions (inconsistent)
2. One unique solution (consistent, square or tall matrices)
3. Infinitely many solutions (consistent, underdetermined, square or wide matrices)



These possibilities can be check by considering $r := \operatorname{rank}(A)$:

- $r < r_b$: case 1, no solutions.
- $r = r_b$: case 1 or 2, at least one solution.
    - $r_b = n$: case 2, a unique solution.
    - $r_b < n$: case 3, infinitely many solutions.

1.5.2 Determinants

Proposition 1.5.4 ***(?).***

$$\det (A \mod p) \mod p \equiv (\det A) \mod p$$

Proposition 1.5.5 ***(Inverse of a 2×2 matrix).***
For 2×2 matrices,

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

In words, swap the main diagonal entries, and flip the signs on the off-diagonal.

Proposition 1.5.6 ***(Properties of the determinant).***
Let $A \in \operatorname{Mat}(m, n)$, then there is a function

$$\begin{aligned} \det : \operatorname{Mat}(m, m) &\to \mathbb{R} \\ A &\mapsto \det(A) \end{aligned}$$

satisfying the following properties:

- det is a group homomorphism onto $(\mathbb{R}, \cdot)$:

$$\det(AB) = \det(A)\det(B)$$

  - Some corollaries:

$$\det A^k = k \det A$$
$$\det(A^{-1}) = (\det A)^{-1} \det(A^t) \qquad\qquad = \det(A).$$

- Invariance under adding scalar multiples of any row to another:

$$\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \end{bmatrix} = \det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i + t\mathbf{a_j} & — \\ & \vdots & \end{bmatrix}$$

- Sign change under row permutation:

$$\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \\ — & \mathbf{a}_j & — \\ & \vdots & \end{bmatrix} = (-1)\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_j & — \\ & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \end{bmatrix}$$

  - More generally, for a permutation $\sigma \in S_n$,

$$\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \\ — & \mathbf{a}_j & — \\ & \vdots & \end{bmatrix} = (-1)^{\operatorname{sgn}(\sigma)}\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_{\sigma(j)} & — \\ & \vdots & \\ — & \mathbf{a}_{\sigma(i)} & — \\ & \vdots & \end{bmatrix}$$

- Multilinearity in rows:

$$\det \begin{bmatrix} & \vdots & \\ — & t\mathbf{a}_i & — \\ & \vdots & \end{bmatrix} = t\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \end{bmatrix}$$
$$\det \begin{bmatrix} — & t\mathbf{a}_1 & — \\ — & t\mathbf{a}_2 & — \\ & \vdots & \\ — & t\mathbf{a}_m & — \end{bmatrix} = t^m \det \begin{bmatrix} — & \mathbf{a}_1 & — \\ — & \mathbf{a}_2 & — \\ & \vdots & \\ — & \mathbf{a}_m & — \end{bmatrix}$$
$$\det \begin{bmatrix} — & t_1\mathbf{a}_1 & — \\ — & t_2\mathbf{a}_2 & — \\ & \vdots & \\ — & t_m\mathbf{a}_m & — \end{bmatrix} = \prod_{i=1}^{m} t_i \det \begin{bmatrix} — & \mathbf{a}_1 & — \\ — & \mathbf{a}_2 & — \\ & \vdots & \\ — & \mathbf{a}_m & — \end{bmatrix}.$$

- Linearity in each row:

$$\det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i + \mathbf{a}_j & — \\ & \vdots & \end{bmatrix} = \det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_i & — \\ & \vdots & \end{bmatrix} + \det \begin{bmatrix} & \vdots & \\ — & \mathbf{a}_j & — \\ & \vdots & \end{bmatrix}.$$

- $\det(A)$ is the volume of the parallelepiped spanned by the columns of A.
- If any row of A is all zeros, $\det(A) = 0$.

Proposition 1.5.7 ***(Characterizing singular matrices).***
TFAE:

- $\det(A) = 0$
- A is singular.

1.5.3 Computing Determinants

Useful shortcuts:

- If A is upper or lower triangular, $\det(A) = \prod_i a_{ii}$.

Definition 1.5.6 (Minors)
The **minor** M_{ij} of $A \in \mathrm{Mat}(n, n)$ is the *determinant* of the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column from A.

Definition 1.5.7 (Cofactors)
The **cofactor** C_{ij} is the scalar defined by

$$C_{ij} := (-1)^{i+j} M_{ij}.$$

Proposition 1.5.8 ***(Laplace/Cofactor Expansion).***
For any fixed i, there is a formula

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}.$$

Example 1.5.1(?): Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Then

$$\det A = 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = 1 \cdot (-3) - 2 \cdot (-6) + 3 \cdot (-3) = 0.$$

Proposition 1.5.9 ***(Computing determinant from RREF).***
$\det(A)$ can be computed by reducing A to $\mathrm{RREF}(A)$ (which is upper triangular) and keeping track of the following effects:

- $R_i \leftarrow R_i \pm tR_j$: no effect.
- $R_i \iff R_j$: multiply by (-1).
- $R_i \leftarrow tR_i$: multiply by t.

1.5.4 Inverting a Matrix

Proposition 1.5.10 ***(Cramer's Rule).***
Given a linear system $A\mathbf{x} = \mathbf{b}$, writing $\mathbf{x} = [x_1, \cdots, x_n]$, there is a formula

$$x_i = \frac{\det(B_i)}{\det(A)}$$

where B_i is A with the ith column deleted and replaced by $\mathbf{b}$.

Proposition 1.5.11 ***(Gauss-Jordan Method for inverting a matrix).***
Under the equivalence relation of elementary row operations, there is an equivalence of augmented matrices:

$$\left[A \mid I\right] \sim \left[I \mid A^{-1}\right]$$

where I is the $n \times n$ identity matrix.

Proposition 1.5.12 ***(Cofactor formula for inverse).***

$$A^{-1} = \frac{1}{\det(A)} [C_{ij}]^t.$$

where C_{ij} is the *cofactor*(Definition 1.5.7) at position i, j.[a]

[a] Note that the matrix appearing here is sometimes called the *adjugate*.

Example 1.5.2(Inverting a 2×2 matrix):

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad \text{where } ad-bc \neq 0$$

What's the pattern?

1. Always divide by determinant
2. Swap the diagonals
3. Hadamard product with checkerboard

Example 1.5.3(Inverting a 3×3 matrix):

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix}$$

$$A^{-1} := \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{-1} = \frac{1}{\det A}\begin{bmatrix} ei-fh & -(bi-ch) & bf-ce \\ -(di-fg) & ai-cg & -(af-cd) \\ dh-eg & -(ah-bg) & ae-bd \end{bmatrix}.$$

The pattern:

1. Divide by determinant
2. Each entry is determinant of submatrix of A with corresponding col/row deleted
3. Hadamard product with checkerboard

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

4. Transpose at the end!!

1.5.5 Bases for Spaces of a Matrix

Let $A \in \mathrm{Mat}(m,n)$ represent a map $T : \mathbb{R}^n \to \mathbb{R}^m$.

Add examples.

Definition 1.5.8 (Pivot)
?

todo

Proposition 1.5.13.

$$\dim \operatorname{rowspace}(A) = \dim \operatorname{colspace}(A).$$

The row space

$$\operatorname{im}(T)^{\vee} = \operatorname{rowspace}(A) \subset \mathbb{R}^n.$$

Reduce to RREF, and take nonzero rows of RREF(A).

The column space

$$\operatorname{im}(T) = \operatorname{colspace}(A) \subseteq \mathbb{R}^m$$

Reduce to RREF, and take columns with pivots from original A.

Remark 1.5.4: Not enough pivots implies columns don't span the entire target domain

The nullspace

$$\ker(T) = \operatorname{nullspace}(A) \subseteq \mathbb{R}^n$$

Reduce to RREF, zero rows are free variables, convert back to equations and pull free variables out as scalar multipliers.

Eigenspaces For each $\lambda \in \operatorname{Spec}(A)$, compute a basis for $\ker(A - \lambda I)$.

1.5.6 Eigenvalues and Eigenvectors

Definition 1.5.9 (Eigenvalues, eigenvectors, eigenspaces)
A vector $\mathbf{v}$ is said to be an **eigenvector** of A with **eigenvalue** $\lambda \in \operatorname{Spec}(A)$ iff

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

For a fixed λ, the corresponding **eigenspace** E_λ is the span of all such vectors.

Remark 1.5.5:

- Similar matrices have identical eigenvalues and multiplicities.
- Eigenvectors corresponding to distinct eigenvalues are **always** linearly independent
- A has n distinct eigenvalues $\implies$ A has n linearly independent eigenvectors.
- A matrix A is diagonalizable $\iff$ A has n linearly independent eigenvectors.

Proposition 1.5.14 ***(How to find eigenvectors).***
For $\lambda \in \mathrm{Spec}(A)$,

$$\mathbf{v} \in E_\lambda \iff \mathbf{v} \in \ker(A - I\lambda).$$

Remark 1.5.6: Some miscellaneous useful facts:

- $\lambda \in \mathrm{Spec}(A) \implies \lambda^2 \in \mathrm{Spec}(A^2)$ with the same eigenvector.
- $\prod \lambda_i = \det A$
- $\sum \lambda_i = \mathrm{Tr}\ A$

Finding generalized eigenvectors

todo

Diagonalizability

Remark 1.5.7: An $n \times n$ matrix P is diagonalizable iff its eigenspace is all of $\mathbb{R}^n$ (i.e. there are n linearly independent eigenvectors, so they span the space.)

Remark 1.5.8: A is diagonalizable if there is a basis of eigenvectors for the range of P.

1.5.7 Useful Counterexamples

$$A := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \implies A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}, \qquad \mathrm{Spec}(A) = [1, 1]$$

$$A := \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \implies A^2 = I_2, \qquad \mathrm{Spec}(A) = [1, -1]$$