

Linear Algebra Notes

Sam Lowe

September 8, 2021

Contents

1	Linear Systems and Matrices	3
1.1	Gauss-Jordan Reduction	3
1.2	Reduced Row-Echelon Form	3
1.3	LU-Decomposition	3
2	Matrix Operations	3
2.1	Matrix Addition	3
2.2	Matrix Multiplication	3
2.3	The Transpose	3
3	Vectors and Vector Spaces	3
3.1	Subspaces	5
3.2	Linear Dependence and Independence	5
4	Basis	5
4.1	Orthogonal Bases and Orthogonal Projections	5
5	Linear Transformations	5
5.1	Operators	6
5.2	Products and Quotients of Vector Spaces	6
6	Determinants	7
6.1	Properties of the Determinant	7
7	Eigenvalues and Eigenvectors	8

8	Matrix Diagonalization	8
9	Symmetric Matrices	8
9.1	Orthogonal Matrices	9
10	Positive and Negative Definite Matrices	9
11	Singular Value Decomposition	9
12	To Be Sorted	9
12.1	Matrices	9
12.2	Determinants	10
12.3	Linear Independence	11
12.4	Determinant	12
12.5	Consequences of a Zero Determinant	12
12.6	Row Operations as Matrices	12
12.7	Eigenvalues and Eigenvectors	12
12.8	Matrix Diagonalization	13
12.9	Dominant Eigenvectors	14

1 Linear Systems and Matrices

1.1 Gauss-Jordan Reduction

1.2 Reduced Row-Echelon Form

1.3 LU-Decomposition

The purpose of **LU-Decomposition** is to rewrite (or *decompose*) a matrix A as the product of a **L**ower-triangular matrix and an **U**pper-triangular matrix to

2 Matrix Operations

2.1 Matrix Addition

2.2 Matrix Multiplication

Row operations can be thought of as left-multiplication by an invertible matrix.

This matrix can be found by performing the desired row operations on the identity matrix.

Swapping the rows of the identity matrix yields a permutation matrix.

2.3 The Transpose

- $(A^t)^t = A$
- $(A + B)^t = A^t + B^t$
- $(rA)^t = r(A^t)$
- $(AB)^t = A^t B^t$

3 Vectors and Vector Spaces

A **vector space over a field** is a set of objects, called **vectors**, which can be added together and multiplied by (or **scaled**) by an element of the field (that element is called a **scalar**). A vector space over \mathbb{R} is usually called a

real vector space and a vector space over \mathbb{C} is usually called a **complex vector space**. For a vector space V over a field F , the following properties hold:

1. **Additive Closure:** $u + v \in V$ for all $u, v \in V$
2. **Additive Identity:** There exists some $0 \in V$, called the **zero vector**, such that $v + 0 = v$ for all $v \in V$
3. **Additive Inverse:** For each $v \in V$, there exists some $-v \in V$ such that $v + (-v) = 0$
4. **Associativity:** $u + (v + w) = (u + v) + w$ for all $u, v, w \in V$
5. **Commutativity:** $u + v = v + u$ for all $u, v \in V$
6. **Multiplicative Closure:** $av \in V$ for all $a \in F$ and all $v \in V$
7. **Compatibility:** $a(bv) = (ab)v$ for all $a, b \in F$ and all $v \in V$
8. **Multiplicative Identity:** $1v = v$ for all $v \in V$, where 1 is the multiplicative identity of F
9. **Distributivity Over Vector Addition:** $a(u + v) = au + av$ for all $a \in F$ and all $u, v \in V$
10. **Distributivity Over Field Addition:** $(a + b)v = av + bv$ for all $a, b \in F$ and all $v \in V$

The first five properties show that V is an Abelian group under addition.

The term vector is used in a broad sense to describe objects that can be added to each other and scaled by some field F . Some common examples of vectors are lists or tuples of numbers, lines in space, the complex numbers, matrices, and functions from any fixed set S to a field F . Polynomials, continuous functions, integrable functions, and differentiable functions all form vector spaces. The solution space to a homogeneous linear differential equation (for example $y' = y$ or $y'' + 5y + 6y = 0$) is also a vector space. A vector v is often denoted \vec{v} in certain contexts.

3.1 Subspaces

A **subspace**, **linear subspace**, or **vector subspace** is a vector space that is a subset of some larger space. For example, if V is a vector space over F and $W \subset V$ is vector space over F under the operations of V , then W is a **subspace of V** . All vector spaces have at least two subspaces, called the **trivial subspaces**: the vector space itself, and the **zero space** $= \{\vec{0}\}$.

Definition Direct Sum

3.2 Linear Dependence and Independence

4 Basis

A **basis** for a vector space is a linearly independent set of vectors that spans that space. The number of vectors is always the same as the — of the vector space that it spans.

4.1 Orthogonal Bases and Orthogonal Projections

5 Linear Transformations

A **linear transformation** (or **linear map**)

Any $m \times n$ matrix A represents a linear transformation $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Even when A is a real matrix, the roots of its characteristic polynomial could be complex. If a given root λ is complex, its associated eigenvectors are complex and therefore doesn't qualify as eigenvectors for $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$; consequently, λ doesn't qualify as an eigenvalue.

However, if we allow complex solutions by consider A to be a complex transformation $A : \mathbb{C}^n \rightarrow \mathbb{C}^n$, then a complex root λ is an eigenvalue because complex solutions are now valid eigenvectors.

When A 's characteristic polynomial only has real coefficients, its complex roots come in complex conjugate pairs: $p(\lambda) = 0$ and $p(\bar{\lambda}) = 0$. Similarly, if v is an eigenvector for λ , then \bar{v} must also be an eigenvector because

$$Av = \lambda v \leftrightarrow \overline{Av} = \overline{\lambda v} = \bar{\lambda} \bar{v}$$

So we conclude that when A is a square matrix with real entries, then (λ, v) is an eigenpair if and only if $(\bar{\lambda}, \bar{v})$.

5.1 Operators

A linear map from a space V to itself is called an **operator**. The notation $\mathcal{L}(V) = \mathcal{L}(V, V)$ denotes the set of all operators on V .

For finite-dimensional vector spaces, a map is injective if and only if it is surjective.

5.2 Products and Quotients of Vector Spaces

Suppose V_1, \dots, V_m are vector spaces over F . Their product $V_1 \times \dots \times V_m$ is defined by $\{(v_1, \dots, v_m) : v_1 \in V_1, \dots, v_m \in V_m\}$.

- Addition on $V_1 \times \dots \times V_m$ is defined by

$$(v_1, \dots, v_m) + (u_1, \dots, u_m) = (v_1 + u_1, \dots, v_m + u_m)$$

- Scalar multiplication on $V_1 \times \dots \times V_m$ is defined by

$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m)$$

Clearly $V_1 \times \dots \times V_m$ is a vector space.

Theorem The dimension of a product is the sum of dimensions. Suppose V_1, \dots, V_m are finite-dimensional vector spaces. Then $V_1 \times \dots \times V_m$ is finite-dimensional with $\dim(V_1 \times \dots \times V_m) = \dim(V_1) + \dots + \dim(V_m)$.

Proof Consider the basis of each vector space. For each basis vector of each V_j , consider the element in $V_1 \times \dots \times V_m$ that consists of only that basis vector in the j th slot and contains 0 (the appropriate additive identity element) in each other slot. The list of all such elements of $V_1 \times \dots \times V_m$ is linearly independent and spans the space, and is thus a basis of length $\dim(V_1) + \dots + \dim(V_m)$ as desired.

Suppose that U_1, \dots, U_m are subspaces of V . Define a linear map $\Gamma : U_1 \times \dots \times U_m \rightarrow U_1 + \dots + U_m$ by

$$\Gamma(u_1, \dots, u_m) = u_1 + \dots + u_m$$

Then $U_1 + \dots + U_m$ is a direct sum if and only if Γ is injective (equivalently, invertible).

A sum is a direct

6 Determinants

One interpretation of an $n \times n$ matrix is as a parallelogram in \mathbb{R}^n with one of its vertices resting at the origin. The **determinant** of a matrix represents how much area (or volume) that matrix takes up. The determinant is only defined for square matrices.

The formula for the determinant of a 2×2 matrix is

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

and the formula for a 3×3 matrix is

$$\begin{aligned} |A| &= \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \end{aligned}$$

Each determinant of a 2×2 matrix in this equation is called a **minor** of the matrix A . The recursive extension of this formula for the determinant of an $n \times n$ matrix is known as **Laplace expansion**.

6.1 Properties of the Determinant

For $n \times n$ matrices A and B ,

- $\det(AB) = \det(A)\det(B)$
- $\det(A^T) = \det(A)$
- $\det(A^{-1}) = \det(A)^{-1} = \frac{1}{\det(A)}$
- $\det(cA) = c^n \det(A)$
- If A is triangular, then its determinant equals the product of its diagonal entries
- If any row or column in A has all elements equal to zero, then $\det(A) = 0$
- If the rows or columns of A are not linearly independent, $\det(A) = 0$

7 Eigenvalues and Eigenvectors

8 Matrix Diagonalization

An $n \times n$ matrix A is called **diagonalizable** if it has n linearly independent eigenvectors, in which case we call those eigenvectors an **eigenbasis**.

If v_1, \dots, v_k are eigenvectors for A corresponding to distinct eigenvalues, then v_1, \dots, v_k are linearly independent. (Worldwide Differential Equations, p.188)

If $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable. However, it is not necessary for A to have distinct eigenvalues in order for it to be diagonalizable.

A square matrix A is diagonalizable if and only if the algebraic multiplicity $m_a(\lambda)$ and the geometric multiplicity $m_g(\lambda)$ are equal for each eigenvalue λ .

These above results apply for both complex and real eigenvalues and eigenvectors.

A square matrix has a basis of eigenvectors if and only if it is similar to a diagonal matrix.

Similar matrices have the same eigenvalues and geometric multiplicities.

9 Symmetric Matrices

Recall that a square matrix is symmetric if $A^T = A$. Symmetric matrices are always diagonalizable.

(Worldwide Differential Equations, Section 6.3 Page 193)

If A is a real symmetric $n \times n$ matrix, then:

- It is diagonalizable,
- All eigenvalues of A are real,
- Eigenvectors corresponding to distinct eigenvalues are orthogonal,
- A has a set of n orthonormal vectors.

Since a real symmetric matrix A has all real eigenvalues, we can also take eigenvectors to be real. Consequently, we use \mathbb{R}^n as our vector space and use the dot product to be the inner product of that space.

9.1 Orthogonal Matrices

A square real matrix O is called orthogonal if $O^T = O^{-1}$.

A real $n \times n$ matrix O is orthogonal if and only if its column vectors form an orthonormal basis for \mathbb{R}^n .

A real $n \times n$ matrix O is orthogonal if and only if $|Ov| = |v|$ for all $v \in \mathbb{R}^n$.
Proof:

$$|Ov|^2 = Ov \cdot Ov = O^T Ov \cdot v = O^{-1} Ov \cdot v = v \cdot v = |v|^2$$

If O is an orthogonal $n \times n$ matrix and v, w are orthogonal vectors in \mathbb{R}^n , then Ov and OW are orthogonal vectors in \mathbb{R}^n . Proof (p 197)

...

10 Positive and Negative Definite Matrices

11 Singular Value Decomposition

12 To Be Sorted

The *span* of a set S of vectors, denoted $span(S)$, is the set of all linear combinations of those vectors. $span(S)$ forms vector space.

12.1 Matrices

The entries of the product matrix P of an $l \times m$ matrix A and a $m \times n$ matrix B are computed by multiplying the rows of A by the columns of B .
 $p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}$.

The matrix

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

represents a θ -radian counterclockwise turn about the origin in \mathbb{R}^2 . The clockwise rotation matrix is given by

$$R(-\theta) = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

This group of two-dimensional rotation matrices is commutative under multiplication; higher dimensional rotation matrices are not commutative.

The *row space* of a matrix A is the span of the row vectors of A . The dimension of the row space is the number of independent rows.

The nullspace of a matrix A , denoted $N(A)$, is the space of all vectors v such that $Av = \vec{0}$. The nullspace is perpendicular to the row space, because for any row vector of A , the dot product of that row with v will equal 0.

The column space of a matrix A , denoted $C(A)$, is the span of the column vectors of A . $C(A^T)$ is the row space. The column space and row space of A have the same dimension; that is to say that the number of independent rows that a matrix has is equal to the number of independent columns that the matrix has.

The left nullspace of a matrix A , denoted $N(A^T)$, is the space of all vectors w such that $Aw = \vec{0}$. The left nullspace is perpendicular to the column space, because for any column vector of A , the dot product of that row with v will equal 0.

Theorem. (Rank-Nullity Theorem)

Let V, W be vector spaces, where V is finite dimensional. Let $T : V \rightarrow W$ be a linear transformation. Then $Rank(T) + Nullity(T) = dim(V)$.

12.2 Determinants

A square matrix's rows are linearly independent if and only if its columns are also linearly independent. A square matrix's rows are linearly dependent if and only if its columns are also linearly dependent.

The determinant of a triangular matrix is the product of the entries in its main diagonal. (Try computing it — every term beyond the first is guaranteed to be zero! Alternately, eliminate the non-leading entries, calculate the determinant, and then simplify.)

For an $n \times n$ diagonal matrix A raised to a power m ,

$$\begin{bmatrix} a_{1,1} & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{bmatrix}^m = \begin{bmatrix} a_{1,1}^m & 0 & \cdots & 0 \\ 0 & a_{2,2}^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n}^m \end{bmatrix}$$

For invertible $n \times n$ matrices A_0, A_1, \dots, A_n , the product $A_0 A_1 \cdots A_n$ is invertible. Additionally, $A_0^{-1} A_1^{-1} \cdots A_n^{-1} = (A_0 A_1 \cdots A_n)^{-1}$.

$$\det(A_1 A_2) = \det(A_1) \det(A_2)$$

The inverse of matrix is unique; $AA^{-1} = A^{-1}A = I$.

If A, B are $n \times n$ square matrices and $AB = I$, then A and B are invertible and $A = B^{-1}$.

Invertible Matrix Theorem

Let A be an $n \times n$ square matrix, X_1 be a constant $n \times 1$ column matrix, and Z_1 be the $n \times 1$ zero matrix. The following statements are equivalent:

- A is invertible.
- The reduced row-echelon form of $A = I$.
- The unique solution to $B = AX_1$ is $X_1 = A^{-1}B$.
- The unique solution to $AX_1 = Z_1$ is $X_1 = Z^1$.

Definition: Elementary An $n \times n$ matrix E is *elementary* if it is the result of applying one of the three Gaussian row operations to the $n \times n$ identity matrix.

The matrices which result from swapping the rows of the identity matrix are called *permutation matrices*.

The matrix which results from performing a row operation on a matrix M is the same as the matrix which results from the product EM , where E is the elementary matrix constructed by using that row operation on the identity matrix.

12.3 Linear Independence

A set of vectors can be tested for linearly independence by putting the associated matrix into reduced row echelon form. If each variable x_1, x_2, \dots, x_n equals zero, then the system is linearly independent. If at least one of the variables is nonzero, the system is linearly dependent.

A set of n $n \times 1$ vectors can be shown to be independent if the determinant of the associated matrix is nonzero (equivalently, if the matrix's reduced row-echelon form is the identity matrix); if the determinant is zero (the matrix's reduced row-echelon form is not the identity matrix), then the vectors are linearly dependent. (This is because the column vectors are linearly independent if and only if the row vectors are linearly independent).

12.4 Determinant

The identity matrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

is an $n \times n$ square matrix with all diagonal entries equal to 1 and all other entries equal to 0.

12.5 Consequences of a Zero Determinant

12.6 Row Operations as Matrices

A row operation on a matrix A can be represented as left multiplication of A by an invertible matrix.

12.7 Eigenvalues and Eigenvectors

An eigenvector \vec{v} of an $n \times n$ square matrix A is a nonzero $n \times 1$ column vector satisfying the equation $A\vec{v} = \lambda\vec{v}$ for some scalar $\lambda \in \mathbb{C}$ (which is that eigenvector's associated eigenvalue). In other words, an eigenvector is a vector which is only scaled by its associated eigenvalue λ (and therefore remain on its own span) when left-multiplied by the matrix A .

The equation $A\vec{v} = \lambda\vec{v}$ can be rearranged to yield $(A - \lambda I)\vec{v} = \vec{0}$. If left-multiplied by $(A - \lambda I)^{-1}$, $\vec{v} = \vec{0}$ will be the only solution, so eigenvectors only exist if $A - \lambda I$ is not invertible; in other words, when

$$\det(A - \lambda I) = 0$$

The eigenvalues of A are the zeroes of this polynomial (which is the characteristic equation of A). The same eigenvalue may occur more than once; the *multiplicity* of an eigenvalue is the number of times the eigenvalue appears as a root. Each eigenvalue λ with multiplicity m has a set of up to m linearly independent eigenvectors; the span of this set of eigenvectors is called an *eigenspace*. Every vector in that eigenspace is an eigenvector for λ .

An eigenvalue with multiplicity 1 is sometimes called *simple*.

Eigenvalues can be substituted back into the eigenvalue equation $(A - \lambda I)$ to solve for the values of x_1, x_2, \dots, x_n in the corresponding $n \times 1$ eigenvector.

If A is an $n \times n$ matrix with only real numbers and if $\lambda_1 = a + bi$ is an eigenvalue with eigenvector \vec{v}^1 . Then $\lambda_2 = \overline{\lambda_1} = a - bi$ is also an eigenvalue and its eigenvector is the conjugate of \vec{v}^1 .

The characteristic equation of an $n \times n$ matrix A will not have more than n roots, so A will have no more than n eigenvalues and no more than n eigenvectors. (These eigenvalues and eigenvectors can be duplicates).

12.8 Matrix Diagonalization

The relationship between an $n \times n$ matrix A , its eigenvalues, and its eigenvectors can be written:

$$\begin{aligned} A\vec{v}_1 &= \lambda_1\vec{v}_1 \\ A\vec{v}_2 &= \lambda_2\vec{v}_2 \\ &\vdots \\ A\vec{v}_n &= \lambda_n\vec{v}_n \end{aligned}$$

or equivalently in matrix form

$$A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Letting P be the matrix of eigenvectors and D be the matrix of eigenvalues, we can rewrite the above equation as

$$A = PDP^{-1}$$

By raising both sides of this equation to some power arbitrary power m and canceling values of P and P^{-1} , we see

$$A^m = PD^mP^{-1}$$

It should be noted that there are other possible values of P and D . However, this algorithm ensures that D is diagonal, so we can exploit the fact

that

$$D^m = \begin{bmatrix} \lambda_1^m & 0 & \cdots & 0 \\ 0 & \lambda_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^m \end{bmatrix}$$

to find values for $\sin D$, e^D , and other functions that can be expressed via power series (and therefore calculate $\sin A$, e^A , etc).

12.9 Dominant Eigenvectors

An eigenvector is said to be *dominant* if its eigenvalue is greater than the eigenvalues of the other eigenvectors.

Consider a diagonalizable matrix A with real-valued eigenvectors. Repeatedly multiplying a vector v by A will tend to “pull” the product vectors towards A ’s dominant eigenvector. This happens because the product vector can be expressed as a linear combination of A ’s eigenvectors, and the dominant eigenvector will be increased the most (and therefore will grow faster and hold more weight in the final linear combination than the other eigenvectors).