

Linalg Fall 2023 Notes

Samuel Lindskog

December 13, 2023

Vectors, Matrices, and Linear Systems

Definition 1.1 (Elementary Row Operations). There are three elementary row operations:¹

1. (Row Interchange) Interchange the i th and j th row vectors in a matrix.
2. (Row Scaling) Multiply the i th row vector in a matrix by a nonzero scalar s .
3. (Row Addition) Add to the i th row vector of a matrix s times the j th row vector.

Definition 1.2 (Elementary Matrix). A matrix which differs from the identity matrix by one elementary row operation.

Definition 1.3 (Transpose Matrix). The matrix B is the transpose of matrix A if $\forall b \in B \wedge \forall a \in A$, we have $b_{ij} = a_{ji}$

Definition 1.4 (Symmetric Matrix). If A is a matrix and $A = A^T$, then A is symmetric.²

Definition 1.5 (Inverse Matrix). The inverse of an $n \times n$ matrix A is the $n \times n$ matrix denoted A^{-1} such that $AA^{-1} = I$.³

Definition 1.6 (Row-Echelon Form). A matrix is in row-echelon form if it satisfies two conditions:

1. All rows containing only zeros appear below rows with nonzero entries.
2. The first nonzero entry in any row appears in a column to the right of the first nonzero entry in any preceding row.⁴

Definition 1.7 (Reduced Row-Echelon Form). A matrix in row-echelon form with all pivots equal to 1 and with zeros above as well as below each pivot is said to be in reduced row-echelon form.⁵

Row Echelon Form Calculation Procedures

The procedure below is known as Gauss reduction with back substitution.

1. Use row interchange to ensure the first row has a nonzero entry in the furthest left nonzero column.

¹ If a matrix B can be obtained from a matrix A by means of row operations, then A is *row equivalent* to B . If two matrices are row-equivalent, they have the same solutions sets.

² Symmetric matrices have the property that any two symmetric matrices multiplied by any two scalars, added to each other form a symmetric matrix.

³ $A^{-1}A = I$ as well. To think about why this is, see that when using the gauss-jordan method, the right side of the augmented matrix is the same as the left side after it has been reduced.

⁴ The first nonzero entry in a row is the pivot for that row. The number of pivots in a reduced matrix is equal to the dimension of that matrix

⁵ The reduced row-echelon form of a matrix is unique.

2. Use row scaling to change the pivot of the first row to 1.
3. Use row add the first row to the other rows so that each entry below the first row's pivot is zero.
4. Mentally cross off first row and column, and repeat with new matrix from step 1 until all pivots are 1.

Definition 1.8 (Consistency). A linear system having no solutions is inconsistent. If it has one or more solutions, the linear system is said to be consistent.

Definition 1.9 (Matrix Multiplication). Multiplication between two matrices is defined as follows:

$$EA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

Definition 1.10 (Homogeneous System). A linear system $Ax = b$ is homogeneous if $b = 0$.⁶

Definition 1.11 (Row Space). The span of the row vectors of a matrix A is the row space of A .

Definition 1.12 (Column Space). The span of the column vectors of a matrix A is the column space of A .

Definition 1.13 (Nullspace). The solution set of $Ax = 0$, is the nullspace of A .

Theorem 1.1 (Column Space Criterion). A linear system $Ax = b$ has a solution if and only if b is in the column space of A .

Theorem 1.2. Let $Ax = b$ be a linear system. If p is any particular solution of $Ax = b$ and h is a solution of the corresponding homogeneous system $Ax = 0$, then $p + h$ is a solution of $Ax = b$.⁷

⁶ The solution set of every homogenous system with n unknowns is a subspace of \mathbb{R}^n .

⁷ Every solution has this form $p + h$, so that the general solution is $x = p + h$

Dimension, Rank, and Linear Transformations

Definition 1.14 (Ordered Basis). By convention, set notation does not denote order; for example, a basis $\{b_1, b_2\}$ is equal to a basis $\{b_2, b_1\}$. To denote order, we use parentheses, $()$. We denote an *ordered basis* of n vectors in V by $B = (b_1, b_2, \dots, b_n)$.⁸

Definition 1.15 (Standard Ordered Basis). We regard the standard basis vectors as having a natural order, $e_1 = [1, 0]$, $e_2 = [0, 1]$. The ordered basis $(E = e_1, \dots, e_n)$ is the *standard ordered basis* for \mathbb{R}^n .⁹

Corollary 1.1. The coordinate vector of $v + w$ relative to basis B is:

$$(v + w)_B = v_B + w_B$$

⁸ Ordered bases are used to specify the coordinate vector of a vector relative to a basis. Otherwise, how would you know which individual basis vectors the elements of v_B refer to?

⁹ To calculate a coordinate vector for v relative to basis B , use Gauss-Jordan reduction of the form $[b_1, \dots, b_n \mid v]$ to find v_B .

Definition 1.16 (Linear Transformation of Real Vector Space). A function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear transformation if it satisfies two conditions:

1. $T(u + v) = T(u) + T(v)$ (Preservation of addition)
2. $T(ru) = rT(u)$ (Preservation of scalar multiplication)

Matrices and Linear Transformations

$T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $T([x_1, x_2]) = [x_2, x_1 - x_2, 2x_1 + x_2]$ is a linear transformation. This can be shown in column-vector notation as:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_2 \\ x_1 - x_2 \\ 2x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

If $A = [T(e_1) \ \dots \ T(e_n)]$ then $T(x) = Ax$ for each column vector $x \in \mathbb{R}^n$. A is referred to as the *standard matrix representation* of the linear transformation T .

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation with standard matrix representation A .

1. The *range* of $T[\mathbb{R}^n]$ of T is the column space of A in \mathbb{R}^m .¹⁰
2. The *kernel* of T is the nullspace of A and is denoted $\ker(T)$.¹¹
3. If W is a subspace of \mathbb{R}^n , then $T[W]$ is a subspace of \mathbb{R}^m .

¹⁰ The dimension of range T is called the *rank* of T .

¹¹ The dimension of $\ker(T)$ is called the *nullity* of T .

Definition 1.17 (Linear Transformation). A function T that maps a vector space V into a vector space V' is a *linear transformation* if it satisfies two conditions:

1. $T(u + v) = T(u) + T(v)$
2. $T(ru) = rT(u)$

Definition 1.18 (Image). For the linear transformation $T : V \rightarrow V'$, if $W \subseteq V$, then $T[W] = \{T(w) \mid w \in W\}$ is the image of W under T .

Definition 1.19 (Domain, Codomain, and Range). For a linear transformation $T : V \rightarrow V'$, the set V is the *domain* of T , and the set V' is the *codomain* of T . $T[V]$ is the *range* of T .

Definition 1.20 (Inverse Image and Kernel). Suppose a linear transformation $T : V \rightarrow V'$. If $W' \subseteq V'$, then $T^{-1}[W'] = \{v \in V \mid T(v) \in W'\}$ is the *inverse image* of W' under T . In particular, $T^{-1}[\{0'\}]$ is the *kernel* of T , denoted by $\ker(T)$.

Definition 1.21 (Invertible Transformation). Let V and V' be vector spaces. A linear transformation $T : V \rightarrow V'$ is *invertible* if there exists a linear transformation $T^{-1} : V' \rightarrow V$ such that $T^{-1} \circ T$ is the identity transformation on V and $T \circ T^{-1}$ is the identity transformation of V' .

Corollary 1.2. An invertible linear transformation T must be bijective:

1. If $v_1 \neq v_2$, then $T(v_1) \neq T(v_2)$. (injective)¹²
2. If v' is in V' , then $T(v) = v'$ for some $v \in V$. (surjective)

¹² A consequence of this is that the $\ker(T) = \{0\}$

Definition 1.22 (Isomorphism). An isomorphism is a linear transformation $T : V \rightarrow V'$ that is bijective.

Theorem 1.3. Let V be a finite-dimensional vector space with ordered basis $B = (b_1, b_2, \dots, b_n)$. The map $T : V \rightarrow \mathbb{R}^n$ defined by $T(v) = v_B$, the coordinate vector v relative to B , is an isomorphism.

Determinants, and Eigenstuff

Definition 1.23 (Second-order Determinant). The area formed by the parallelogram of two vectors $[a_1, a_2]$ and b_1, b_2 in a 2×2 matrix is called the second-order determinant, and is calculated as follows:

$$\text{Area} = \det(A) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

Definition 1.24 (Third-order Determinant). The volume formed by the parallelepiped of three vectors $[a_1, a_2, a_3]$, $[b_1, b_2, b_3]$ and $[c_1, c_2, c_3]$ is the third-order determinant of these vectors. $(b \times c) \cdot a$ is the determinant of the matrix:

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

and is denoted by $\det(A)$.

Theorem 1.4 (Determinant Criterion for Invertibility). A square matrix A is invertible iff $\det(A) \neq 0$. Equivalently, A is singular iff $\det(A) = 0$.

Corollary 1.3 (Computation of determinants). The determinant of an $n \times n$ matrix A can be computed as follows:

1. Reduce A to an echelon form, using only row addition and row interchanges.
2. If any of the matrices appearing in the reduction contains a row of zeros, then $\det(A) = 0$

3. Otherwise, $\det(A) = (-1)^r \cdot (\text{product of pivots})$, where r is the number of row interchanges performed.

Definition 1.25 (Minor Matrix). The minor matrix A_{ij} of an $n \times n$ matrix $A = [a_{ij}]$ is the $(n-1) \times (n-1)$ matrix obtained by crossing out the i th and j th column of A . Using $|A_{ij}|$ as notation for the determinant of the minor matrix A_{ij} , we can express the determinant of a 3×3 matrix A as:

$$\det(A) = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}|$$

Definition 1.26 (Cofactor). The cofactor of a'_{ij} in A is:

$$a'_{ij} = (-1)^{i+j} \det(A_{ij})$$

Corollary 1.4. The determinant of a 1×1 matrix is its sole entry; it is a first-order determinant. Let $n \geq 1$, and assume that determinants of order less than n have been defined. Let $A = [a_{ij}]$ be an $n \times n$ matrix. The determinant of A is:¹³

$$\det(A) = \sum_{i=1}^n a_{1i} a'_{1i}$$

¹³ A consequence of this is that $\det(A) = \det(A^T)$.

Definition 1.27 (Row-Interchange Property). If two different rows of a square matrix A are interchanged, the determinant of the resulting matrix is $-\det(A)$.

Definition 1.28 (Scalar-Multiplication Property). If a single row of a square matrix A is multiplied by a scalar r , the determinant of the resulting matrix is $r \cdot \det(A)$.

Theorem 1.5 (Determinant Criterion for Invertibility). A square matrix A is invertible iff $\det(A) \neq 0$.

Definition 1.29 (Multiplicative Property). If A and B are $n \times n$ matrices, then $\det(AB) = \det(A) \cdot \det(B)$.

Definition 1.30 (Row-Addition Property). If the product of one row of a square matrix A by a scalar is added to a different row of A , the determinant of the resulting matrix is the same as $\det(A)$.

Definition 1.31 (Eigenvalues and Eigenvectors). Let A be an $n \times n$ matrix. A scalar λ is an *eigenvalue* of A if there is a nonzero column vector v in \mathbb{R}^n such that $Av = \lambda v$. The vector v is then an *eigenvector* of A corresponding to λ .

uses for eigenvectors and eigenvalues

Eigenvectors are good for calculating recursive transformations, because $A^k v = \lambda^k v$. For many matrices A , the computation of

$A^k x$ for a general vector x is greatly simplified by first finding all eigenvalues and eigenvectors of A . Suppose that we can find a basis $\{b_1, \dots, b_n\}$ for \mathbb{R}^n consisting of eigenvectors of A , so that for eigenvalues $\lambda_1, \dots, \lambda_n$ we have:

$$Ab_i = \lambda_i b_i \text{ for } 1 \leq i \leq n$$

To compute A^k , we first express x as a linear combination of these basis eigenvectors:

$$x = d_1 b_1 + \dots + d_n b_n$$

Because $A^k b_i = \lambda_i^k b_i$, we obtain:

$$A^k x = \sum_{i=1}^n \lambda_i^k d_i b_i$$

Theorem 1.6. Let A be an $n \times n$ matrix.

1. If λ is an eigenvalue of A with v as a corresponding eigenvector, then λ^k is an eigenvalue of A^k , again with v as a corresponding eigenvector, for any positive integer k .
2. If λ is an eigenvalue of an invertible matrix A with v as a corresponding eigenvector, then $\lambda \neq 0$ and $1/\lambda$ is an eigenvalue of A^{-1} , again with v as a corresponding eigenvector.
3. If λ is an eigenvalue of A , then the set E_λ consisting of the zero vector together with all eigenvectors of A for this eigenvalue λ is a subspace of n -space, the *eigenspace* of λ .

Theorem 1.7. Let A be an $n \times n$ matrix, and v_1, \dots, v_n be nonzero vectors in n -space. Let C be the $n \times n$ matrix having v_j as its j th column vector, and let D be a diagonal matrix with $\lambda_1, \dots, \lambda_n$ be its scalars from the (1)st to the n th columns respectively. Then $AC = CD$ iff $\lambda_1, \dots, \lambda_n$ are eigenvalues of A and v_j is an eigenvector of A corresponding to λ_j for $j = 1, \dots, n$.

Definition 1.32 (diagonalizable). An $n \times n$ matrix A is *diagonalizable* if there exists an invertible matrix C such that $C^{-1}AC = D$, a diagonal matrix. The matrix C is said to *diagonalize* the matrix A .

Corollary 1.5. An $n \times n$ matrix A is diagonalizable if and only if n -space has a basis consisting of eigenvectors of A .¹⁴

Corollary 1.6. Let an $n \times n$ matrix A have n eigenvectors and eigenvalues, giving rise to the matrices C and D so that $AC = CD$ as described above. If the n eigenvectors are independent, then C is an invertible matrix and $C^{-1}AC = D$. Under these conditions, we have:

$$A^k = CD^k C^{-1}$$

¹⁴ This make sense because if two eigenvectors were linearly dependent, they would both correspond to the same eigenvalue. Linearly independent eigenvectors can occur in the same eigenspace.

Definition 1.33. An $n \times n$ matrix P is *similar* to an $n \times n$ matrix Q if there exists an invertible $n \times n$ matrix C such that $C^{-1}PC = Q$.

Definition 1.34 (Algebraic and Geometric Multiplicity). The *algebraic multiplicity* of an eigenvalue λ_i of A is its multiplicity as a root of the characteristic equation of A . Its *geometric multiplicity* is the dimension of the eigenspace E_λ .

Theorem 1.8. An $n \times n$ matrix A is diagonalizable iff the algebraic multiplicity of each (possible complex) eigenvalue is equal to its geometric multiplicity.

Projection

Theorem 1.9. The projection p of b on $\text{sp}(a)$ in \mathbb{R}^n is:

$$p = \frac{b \cdot a}{a \cdot a} a$$

Definition 1.35 (Orthogonal Complement). Let W be a subspace of \mathbb{R}^n . The set of all vectors in \mathbb{R}^n that are orthogonal to every vector in W is the *orthogonal complement* of W , and is denoted W^\perp . The orthogonal complement of the row space of A , the generating set for W , is the nullspace of A .

Theorem 1.10 (Orthogonal Bases). Let $\{v_1, \dots, v_n\}$ be an orthogonal set of nonzero vectors in \mathbb{R}^n . Then this set is independent and consequently is a basis for the subspace $\text{sp}(v_1, \dots, v_n)$.