Chapter 3 Notes - LA

John Yang

February 24, 2022

Contents

3 Matrices

3.1 Matrix Operations

- A matrix is defined as a rectangular array of numbers called the entries, or elements, of the matrix
- The size of a matrix is based on the number of rows and columns; a matrix with m rows and n columns is an $m \times n$ matrix (m by n).
- Entries of matrix are refered to with double subscripts
- If $m = n$, the matrix is a square matrix. If all nondiagonal entries are 0, the matrix is a diagonal matrix. A diagonal matrix whose diagonal entries are the same is called a scalar matrix. If the scalar on the diagonal is 1 it is an identity matrix.
- Adding matrices: only matrices with the same dimensions can be added. Add each corresponding entry.
- A matrix whose entries are all 0 is a zero matrix denoted by O.
- Multiplying matrices: If A is an $m \times n$ matrix and B is an $n \times r$ matrix, then the product $C = AB$ is an $m \times r$ matrix. The (i, j) entry of the product is computed as follows:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

- Note that if A is an m by n matrix and B is an n by r matrix, AB is an m by r matrix and n must be equal to n.
- The product of two matrices is a dot product.
- Theorem: let A be an $m \times n$ matrix, $\mathbf{e}_i$ a $1 \times m$ standard unit vector, and $\mathbf{e}_j$ an $n \times 1$ standard unit vector. Then:
  - $\mathbf{e}_i A$ is the ith row of A and

- $A\mathbf{e}_j$ is the jth column of A.

- Partitioned matrices: you can divide a matrix into submatrices by partitioning it into blocks.
- Matrx powers:
$$A^k = AA \cdots A$$
by k factors
- If A is a square matrix and r and s are nonnegative integers, then
$$A^r A^s = A^{r+s} \qquad (A^r)^s = A^{rs}$$
- The transpose of an $m \times n$ matrix A is the $n \times m$ matrix of A^T obtained by interchanging the rows and columns of A. That is, the ith column of A^T is the ith row of A for all i
- A square matrix A is defined as symmetric if $A^T = A$; that is, if A is equal to its own transpose.
- A square matrix A is symmetric IFF $A_{ij} = A_{ji}$ for all i and j

3.2 Matrix Algebra

- Properties of matrix addition and scalar multiplication
  - $A + B = B + A$
  - $(A + B) + C = A + (B + C)$
  - $A + O = A$
  - $A + (-A) = O$
  - $c(A + B) = cA + cB$
  - $(c + d)A = cA + dA$
  - $c(dA) = (cd)A$
  - $1A = A$
- Linear independence applies to matrices as well
- Properties of matrix multiplication
  - $A(BC) = (AB)C$
  - $A(B + C) = AB + AC$
  - $(A + B)C = AC + BC$
  - $k(AB) = (kA)B = A(kB)$
  - $I_m A = A = AI_n$ if A is $m \times n$
- Properties of the transpose
  - $(A^T)^T = A$
  - $(A + B)^T = A^T + B^T$
  - $(kA)^T = k(A^T)$
  - $(AB)^T = B^T A^T$
  - $(A^r)^T = (A^T)^r$ for all nonnegative integers r
- Transposing a matrix: like flipping it on its side; rows become columns and columns become rows. Order stays the same; left to right, top to bottom
- If A is a square matrix, then $A + A^T$ is a symmetric matrix
- For any matrix A, AA^T and $A^T A$ are symmetric matrices.

3.3 The Inverse of a Matrix

- If A is an $n \times n$ matrix, an inverse of A is an $n \times n$ matrix A' with the property that

$$AA' = I \quad \text{and} \quad A'A = I$$

  where $I = I_n$, the $n \times n$ identity matrix. If A' exists, then A is called invertible.

- If A is an invertible matrix, then its inverse is unique.

- If A is an invertible $n \times n$ matrix, then the system of linear equations given by $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$ for any $\mathbf{b}$ in $\mathbb{R}^n$

- If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then A is invertible if $ad - bc \neq 0$, in which case

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

  If $ad - bc = 0$, then A is not invertible.

- The expression $ad - bc$ is the determinant of A, given by $\det A$

- Properties of invertible matrices:
  - If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

  - If A is an invertible matrix and c is a nonzero scalar, then cA is an invertible matrix and

$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

  - If A and B are invertible matrices of the same size, then AB is invertible, and

$$(AB)^{-1} = B^{-1}A^{-1}$$

  - If A is an invertible matrix, then A^T is invertible and

$$(A^T)^{-1} = (A^{-1})^T$$

  - If A is an invertible matrix, then A^n is invertible for all nonnegative integers n and

$$(A^n)^{-1} = (A^{-1})^n$$

- The inverse of a product of invertible matrices is the product of their inverses in reverse order.

- If A is an invertible matrix and n is a positive integer, then A^{-n} is defined by

$$A^{-n} = (A^{-1})^n = (A^n)^{-1}$$

- An elementary matrix is one that can be obtained by performing an elementary row operation on an identity matrix.

- Let E by the elementary matrix obtained by performing an elementary row operation on I_n. If the same elementary row operation is performed on an $n \times r$ matrix A, the result is the same as the matrix EA.

- Each elementary matrix is invertible, and its inverse is an elementary matrix of the same type.

- The fundamental theorem of invertible matrices: version 1. Let A be an $n \times n$ matrix. The following statements are equivalent:
  - A is invertible.
  - $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The RREF of A is I_n
  - A is a product of elementary matrices.
- Let A be a square matrix. If B is a square matrix such that either $AB = I$ or $BA = 1$, then A is invertible and $B = A^{-1}$
- Let A be a square matrix. If a sequence of elementary row operations reduces A to I, then the same sequence of elementary row operations transforms I into A^{-1}

3.4 The LU Factorization

- Let A be a square matrix. A factorization of A as $A = LU$, where L is unit lower triangular and U is upper triangular, is called an LU factorization of A.
- If A is a square matrix that can be reduced to REF without using any row interchanges, then A has an LU factorization.
- If A is an invertible matrix that has an LU factorization, then L and U are unique.
- If P is a permutation matrix, then $P^{-1} = P^T$
- Let A be a square matrix. A factorization of A as $A = P^T LU$, where P is a permutation matrix, L is unit lower triangular, and U is upper triangular, is called a $P^T LU$ factorization of A
- Every square matrix has a $P^T LU$ factorization.

3.5 Subspaces, Basis, Dimension, and Rank

- A subspace of $\mathbb{R}^n$ is any collection of S vectors in $\mathbb{R}^n$ such that
  - The zero vector $\mathbf{0}$ is in S
  - If $\mathbf{u}$ and $\mathbf{v}$ are in S, then $\mathbf{u} + \mathbf{v}$ is in S (That is, S is closed under addition).
  - If vbu is in S and c is a scalar, then $c\mathbf{u}$ is in S (S is closed under scalar multiplication).
  - From the previous two conditions, we conclude that S must then be closed under linear combinations: S includes all linear combinations of all vectors $\mathbf{u}_k$ in S.
- Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ be vectors in $\mathbb{R}^n$. Then, $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k)$ is a subspace of $\mathbb{R}^n$
- Let A be an $m \times n$ matrix.
  - The row space of A is the subspace $\text{row}(A)$ of $\mathbb{R}^n$ spanned by the rows of A.
  - The column space of A is the subspace $\text{col}(A)$ of $\mathbb{R}^m$ spanned by the columns of A.
- Let B be any matrix that is row equivalent to a matrix A. Then $\text{row}(B) = \text{row}(A)$
- Let A be an $m \times n$ matrix and let N be the set of solutions to the homogeneous linear system $A\mathbf{x} = \mathbf{0}$. Then N is a subspace of $\mathbb{R}^n$
- Let A be an $m \times n$ matrix. The null space of A is the subspace of $\mathbb{R}^n$ consisting of solutions to the homogeneous linear system $A\mathbf{x} = 0$. It is denoted by $\text{null}(A)$

- Let A be a matrix whose entries are real numbers. For any system of linear equations $A\mathbf{x} = \mathbf{b}$, exactly one of the following is true:
  - There is no solution.
  - There is a unique solution.
  - There are infinitely many solutions.
- A basis for a subspace S of $\mathbb{R}^n$ is a set of vectors in S that
  - spans S and
  - Is linearly independent.
- How to find the bases for the row space, column space, and null space of matrix A:
  - Find the rref R of A.
  - Use the nonzero row vectors of R (containing the leading 1s) to form a basis for row(A)
  - Use the column vectors of A that correspond to the columns of R containing the leading 1s (the pivot columns) to form a basis for col(A)
  - Solve for the leading variables of $R\mathbf{x} = 0$ in terms of the free variables, set the free variables equal to parameters, substitute back into $\mathbf{x}$, and write the result as a linear combination of f vectors (where f is the number of free variables). These f vectors form a basis for null(A)
- The Basis Theorem: Let S be a subspace of $\mathbb{R}^n$. Then any two bases for S have the same number of vectors.
- If S is a subspace of $\mathbb{R}^n$, then the number of vectors in a basis for S is called the dimension of S, denoted dim S.
- The row and column spaces of a matrix A have the same dimension.
- The rank of a matrix A is the dimension of its row and column spaces and is denoted by rank(A).
- For any matrix A,
$$\text{rank}(A^T) = \text{rank}(A)$$
- The nullity of a matrix A is the dimension of its null space and is denoted by nullity(A).
- The Rank Theorem: If A is an $m \times n$ matrix, then
$$\text{rank}(A) + \text{nullity}(A) = n$$
- The fundamental theorem of invertible matrices: version 2. Let A be an $m \times n$ matrix. The following statements are equivalent:
  - A is invertible.
  - $A\mathbf{x} = b$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The rref of A is I_n
  - A is a product of elementary matrices.
  - $\text{rank}(A) = n$
  - $\text{nullity}(A) = 0$
  - The column vectors of A are linearly independent
  - The column vectors of A span $\mathbb{R}^n$.
  - The column vectors of A form a basis for $\mathbb{R}^n$.

- The row vectors of A are linearly independent.
- The row vectors of A span $\mathbb{R}^n$
- The row vectors of A form a basis for $\mathbb{R}^n$.

- Let A be an $m \times n$ matrix. Then
  - $\text{rank}(A^T A) = \text{rank}(A)$
  - The $n \times n$ matrix $A^T A$ is invertible IFF $\text{rank}(A) = n$

- Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be a basis for S. For every vector $\mathbf{v}$ in S, there is exactly one way to write $\mathbf{v}$ as a linear combination of the basis vectors in $\mathcal{B}$:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

- Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be a basis for S. Let $\mathbf{v}$ be a vector in S, and write $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. Then $c_1, c_2, \cdots, c_k$ are called the coordinates of $\mathbf{v}$ with respect to $\mathcal{B}$, and the column vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

  is called the coordinate vector of $\mathbf{v}$ with respect to $\mathcal{B}$.

3.6 Introduction to Linear Transformations

- A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if:
  - $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all u and v in $\mathbb{R}^n$ and
  - $T(c\mathbf{v}) = cT(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$ and all scalars c.

- Let A be an $m \times n$ matrix. Then the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A(\mathbf{x}) = A\mathbf{x} \quad (\text{for } \mathbf{x} \text{ in } \mathbb{R}^n)$$

  is a linear transformation.

- Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is a matrix transformation. More specifically, $T = T_A$, where A is the $m \times n$ matrix

$$A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)]$$

- Let $T : \mathbb{R}^m \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^p$ be linear Transformations. then $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ is a linear transformation whose standard matrices are related by

$$S \circ T = [S][T]$$

- Let s and T be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then S and T are inverse transformations if $S \circ T = I_n$ and $T \circ S = I_n$

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then its standard matrix $[T]$ is an invertible matrix, and

$$[T^{-1}] = [T]^{-1}$$

  ("The matrix of the inverse is the inverse of the matrix").

3.7 Applications

- Markov chain: evolving proccess consisting of a finite number of states.
- Can use linear algebra to analyze probabilities. Consider two-way tables in statistics: working with multiple of these tables as matrices and vectors can allow us to solve probability problems.
- Graphs and digraphs: If G is a graph with n vertices, then its adjacency matrix is the $n \times n$ matrix A (or $A(G)$) defined by

$$a_{ij} = 1 \text{ if there is an edge between vertices i and j, and } a_{ij} = 0 \text{ otherwise.}$$

- If A is the adjacency matrix of a graph G, then the (i, j) entry of A^k is equal to the number of k-paths between vertices i and j.
- If G is a digraph with n vertices, then its adjacency matrix is the $n \times n$ matrix A (or $A(G)$) defined by

$$a_{ij} = 1 \text{ if there is an edge between vertices i and j, and } a_{ij} = 0 \text{ otherwise.}$$

- Error correcting codes: If $k < n$, then any $n \times k$ matrix of the form $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$, where A is an $(n-k) \times k$ matrix over $\mathbb{Z}_2$, is called a standard generator matrix for an (n, k) binary code $T : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$. Any $(n-k) \times n$ matrix of the form $P = [B\ I_{n-k}]$, where B is an $(n-k) \times k$ matrix over $\mathbb{Z}_2$, is called a stardard parity check matrix. The code is said to have length n and dimension k.
- If $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$ is a standard generator matrix and $P = [B\ I_{n-k}]$ is a standard parity check matrix, then P is the parity check matrix associated with G IFF $A = B$. The corresponding n, k binary code is (single) error-correcting IFF the columns of P are nonzero and distinct.
- Summary of error-correcting codes:
  - For $n > k$, and $n \times k$ matrix G and an $(n-k) \times n$ matrix P (with entries in $mathbbZ_2$) are a standard generator matrix and a standard parity check matrix, respectively, for an (n, k) binary code IFF in block form, $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$ and $P = [A\ I_{n-k}]$ for some $(n-k) \times k$ matrix A over $\mathbb{Z}_2$.
  - G encodes a message vector $\mathbf{x}$ in $\mathbb{Z}_2^k$ as a code vector $\mathbf{c}$ in $\mathbb{Z}_2^n$ via $\mathbf{c} = G\mathbf{x}$.
  - G is error-correcting IFF the columns of P are nonzero and distinct. A vector $\mathbf{c}'$ in $\mathbb{Z}_2^n$ is a code vector IFF $P\mathbf{c}' = \mathbf{0}$. In this case, the corresponding message vector is the vector $\mathbf{x}$ in $\mathbb{Z}_2^k$ consisting of the first k components of $\mathbf{c}'$. If $P\mathbf{c}' \neq 0$, then $\mathbf{c}'$ is not a code vector and $P\mathbf{c}'$ is one of the columns of P. If $P\mathbf{c}'$ is the ith column of P, then the error is in the ith component of $\mathbf{c}'$ and we can recover the correct code vector (and hence the message) by changing this component.