Math 307 Review Notes

Reese Critchlow

Midterm 1

LU Decomposition

The gaussian elimination on a matrix A can be expressed using LU Decomposition. LU Decomposition follows the form:

$$A = LU$$

Where L is a unit lower triangular matrix and U is a upper triangular matrix.

Unit Lower Triangular Matricies

A unit lower triangular matrix has the following attributes:

- The matrix is square ($n \times n$).
- The diagonal entries of the matrix are ones and only zeroes are above the ones.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ x & 1 & 0 & 0 \\ x & x & 1 & 0 \\ x & x & x & 1 \end{bmatrix}$$

Upper Triangular Matricies

An upper triangular matrix has the following attribute:

- The matrix has only zeroes below the main diagonal.

$$\begin{bmatrix} 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix}$$

Elementary Row Operations as Matrix Multiplications

The elementary row operations can be expressed as matrix multiplications. They are as follows:

Interchange rows i and j.

Modify the identity matrix such that:

$$a_{i,i} = 0 \qquad a_{j,j} = 0$$
$$a_{i,j} = 1 \qquad a_{j,i} = 1$$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1_i & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1_j & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Multiply a row i by a scalar k.

Modify the identity matrix such that:

$$a_{i,i} = k$$

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & k_i & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Add c times row j to i.

Modify the identity matrix such that:

$$a_{i,j} = c$$

Note: it is required that $i > j$.

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ c_{i,j} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

It is also important to note that the E^{-1} can be acheived by modifying the identity matrix such that:

$$a_{j,i} = -c.$$

Hence, Gaussian elimination can be expressed as $E_3E_2E_1A$.

It is important to note that the order of the inverse is paramount. The inverse of $E_3E_2E_1$ is $E_1^{-1}E_2^{-1}E_3^{-1}$, not $E_3^{-1}E_2^{-1}E_1^{-1}$. In short, inverses are to be applied in the opposite order that the original matricies were applied.

Generally, only the E transformation is used, because the rest of the transformations do not result in a lower triangular matrix.

Theorem: If a matrix A can be converted to row echelon form using only E row operations, then A has an LU decomposition.

Sometimes, the LU decomposition is not attainable, thus, other transformations are permitted, but the LU decomposition will take on the form of $A = PLU$ where P is a permuatation matrix.

Important Properties of the LU Decomposition:

- $\text{rank}(A) = \text{rank}(U)$
- If a is a square matrix, then: $\det(A) = \det(U)$ If A is a square matrix of full rank:
- $\det(A) \neq 0$.
- A is invertible.
- $A\vec{x} = \vec{b}$ has a unique solution.

Rank of a Matrix: As a review, the rank of a matrix is:

- The dimension of the span of the matrix.
- The number of non-zero leading/pivot entries in a matrix which is in row echelon form.

Row Echelon Form: As a review, a matrix is considered to be in row echelon form when

- All rows consisting of only zeroes are at the bottom.
- The pivot entry of any nonzero row is always strictly to the right of the leading coefficient of the row above it.

Inverting a Matrix: To invert a matrix A, create a matix $[A \mid I]$ and use row operations to transform the matrix to a form of $[I \mid A^{-1}]$.

Error Analysis and Matrix Norms

Matrix Norm: The matrix norm or operator norm is defined by:

$$||A|| = \max_{x \neq 0} \frac{||A\vec{x}||}{||\vec{x}||} \qquad \text{where } ||\vec{x}|| \text{is the } e^2 \text{ norm.}$$

The matrix norm describes the maximum stretch of a unit vector. Inverse Matrix Norm: If a matrix A is a square, non-singular matrix, then

$$||A^{-1}|| = \frac{1}{\min_{||\vec{x}||=1} ||A\vec{x}||}.$$

Significance of the Matrix Norm: If a matrix or a vector is obtained emperically, there may be errors in it. Thus, the matrix norm allows us to predict how large the effect of those errors may be.

Condition Number: The condition number of a matrix A is defined as:

$$\text{cond}(A) = ||A|| \cdot ||A^{-1}||$$

Error Bounding: Given a system $A\vec{x} = \vec{b}$, assuming that $A(\vec{x} + \Delta\vec{x}) = \vec{b} + \Delta\vec{b}$, then:

$$\frac{||\Delta\vec{x}||}{||\vec{x}||} \leq \text{cond}(A) \cdot \frac{||\Delta\vec{b}||}{||\vec{b}||}.$$

The equation above describes the error in $\vec{x}$ as a result of errors in $\vec{b}$.

One can also produce an expression for the effects of errors in A:

$$\frac{||\Delta\vec{x}||}{||\vec{x}+\Delta\vec{x}||} \leq \text{cond}(A) \cdot \frac{||\Delta A||}{||A||}$$

A common type of problem is to extract the condition number from the manipulation of an image.
Norm For Diagonal Matricies: The norm of a diagonal matrix D can be obtained with the following formula:

$$D = \max\{|d_k|\}$$

where d_k are the set of diagonal entries in the matrix.

Relations of the Condition Number: It is important to note that:

$$||AB|| \leq ||A|| \cdot ||B|| \qquad \text{thus} \qquad \text{cond}(AB) \leq \text{cond}(A) \cdot \text{cond}(B)$$

Polynomial Interpolation

A polynomial of the form:

$$P_\alpha = \left\{c_0 + c_1 t + c_2 t^2 + \cdots + c_\alpha t^\alpha : c_n \in \mathbb{R}\right\}$$

Can be used to interpolate α number of points, so long as $t_i \neq t_j$.

Given a system with multiple points we can solve a system of equations to obtain the unknonw coefficients, c_n.

Such a system of equations takes the form of $A\vec{c} = \vec{y}$, where A is known as the Vandermonde Matrix.

$$\begin{bmatrix} 1 & t_0 & t_0^2 & \cdots & t_0^\alpha \\ 1 & t_1 & t_1^2 & \cdots & t_1^\alpha \\ 1 & t_2 & t_2^2 & \cdots & t_2^\alpha \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_\alpha & t_\alpha^2 & \cdots & t_\alpha^\alpha \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_\alpha \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_\alpha \end{bmatrix}$$

Determinant of a Vandermonde Matrix: The determinant of a Vandermonde matrix can be obtained by the following formula:

$$\det(A) = \prod_{0 \leq i < j \leq \alpha} (t_j - t_i)$$

It is important to note that in a Vandermonde matrix, the entries $t_0, t_1, \cdots, t_n$ represent time points and the y vector represents their corresponding values.

It is also important to note that the condition number of a Vandermonde matrix gets *very* large when the number of points increases.

Cubic Spline Interpolation

Another method for interpolation is the cubic spline method. Unlike the continuous nature of the Polynomial Interpolation, the cubic spline method is piecewise.

By definition, a cubic spline interpolation has the following features/properties.

- There are N cubic polynomials of the form:
  $p_k(t) = a_k(t - t_{k-1})^3 + b_k(t - t_{k-1})^2 + c_k(t - t_{k-1}) + d_k$
- $p(t)$, $p'(t)$, and $p''(t)$ are continuous
- $p(t_i) = y_i$ for all $i = 0, \cdots, N$

To find a cubic spline interpolation, one must use the coefficient matrix of a cubic spline, which has the following form:

$$C = \begin{bmatrix} a_1 & a_2 & \cdots a_N \\ b_1 & b_2 & \cdots b_N \\ c_1 & c_2 & \cdots c_N \\ d_1 & d_2 & \cdots d_N \end{bmatrix}$$

Hence, it can be concluded from this that every cubic spline interpolation coefficient matrix will have $4N$ terms, where N is the number of data points.

<u>Equations Defining the Cubic Spline:</u> There exist several methods to obtain the equations to solve for a cubic spline. They are as follows:

1. Interpolation
   (a) Left Endpoints
   Basis: $p_k(t_{k-1}) = y_{k-1}$
   Result: $d_k = y_{k-1}$
   Number of Equations: N
   (b) Right Endpoints
   Basis: $p_k(t_k) = y_k$
   Result: $a_k(t_k - t_{k-1})^3 + b_k(t_k - t_{k-1})^2 + c_k(t_k - t_{k-1}) + d_k = y_k$
   Number of Equations: N
2. Continuity of the Derivative
   Basis: $p'_k(t_k) = p'_{k+1}(t_k)$
   Result: $3a_k(t_k - t_{k-1})^2 + 2b_k(t_k - t_{k-1}) + c_k = c_{k+1}$
   Number of Equations: $N - 1$
3. Continuity of the Second Derivative
   Basis: $p''_k(t_k) = p''_{k+1}(t_k)$
   Result: $6a_k(t_k - t_{k-1}) + 2b_k = 2b_{k+1}$
   Number of Equations: $N - 1$
4. Secondary Methods
   (a) Neutral Cubic Spline Conditions Equations: $p''(t_0) = 0$ and $p''_N(t_N) = 0$

<u>Obtaining the Coefficients for a Cubic Spline:</u> First, it is known that $d_n = y_{n-1}$. Thus, the system that needs to be solved is as follows:

$$\begin{bmatrix} A(L_1) & B & 0\cdots 0 & & \\ 0 & A(L_2) & B\cdots 0 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T & 0 & 0 & V & \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ \vdots \\ a_N \\ b_N \\ c_N \end{bmatrix} = \begin{bmatrix} y_1 - y_0 \\ 0 \\ 0 \\ y_2 - y_1 \\ 0 \\ 0 \\ \vdots \\ y_N - y_{N-1} \\ 0 \\ 0 \end{bmatrix}$$

Where $L_k = t_k - t_{k-1}$ and:

$$A(L) = \begin{bmatrix} L^3 & L^2 & L \\ 3L^2 & 2L & 1 \\ 6L & 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -2 & 0 \end{bmatrix} \qquad T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad V = \begin{bmatrix} L_N^3 & L_N^2 & L_N \\ 0 & 0 & 0 \\ 6L_N & 2 & 0 \end{bmatrix}.$$

Subspaces

Definition: A subset $U \subseteq \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ under the following conditions:

1. U contains the zero vector $\vec{0} \in \mathbb{R}$.
2. Closed under addition: $\vec{u_1}, \vec{u_2} \in U \Rightarrow \vec{u_1} + \vec{u_2} \in U$.
3. Closed under scalar multiplication: $\vec{u} \in U, c \in \mathbb{R} \Rightarrow c\vec{u} \in U$.

For example, the smallest subspace of $\mathbb{R}^2$ is $\{\vec{0}\}$, and the largest subspace of $\mathbb{R}^2$ is simply $\mathbb{R}^2$. A common subspace however is any line passing through the origin.

Subspaces of $\mathbb{R}^3$ include lines through the origin and planes containing the origin.

Span

Definition: The span of a set of vectors is the set of all of the possible inear combinations of them.

Determining Span Membership: To determine whether a vector, $\vec{v}$ is within the span of a set of vectors, $\{\vec{u}_n\}$, one can write $\begin{bmatrix} \vec{u_1} & \vec{u_2} & \cdots & \vec{u_n} \mid \vec{v} \end{bmatrix}$ and solve the matrix.

Linear Independence

The vectors $\vec{u_1}, \cdots, \vec{u_m} \in \mathbb{R}$ are said to be linearly independent if $c_1\vec{u_1} + \cdots + c_m\vec{u_m} = \vec{0}$ if and only if the solution is trivial.

A more algorithmic approach to finding linear independence is to solve the matrix $\begin{bmatrix} \vec{u_1} \cdots \vec{u_n} \mid \vec{0} \end{bmatrix}$

Basis and Dimension

Definition: Let $U \subseteq \mathbb{R}^n$ be a subspace. A set of vectors $\{\vec{u_m}\}$ forms a basis of U if:

1. $\{\vec{u_m}\}$ is a linearly independent set.
2. $\text{span}\{\vec{u_m}\} = U$.

Remark: There are infinitley many different bases of a subspace U, but each basis of U has the same number of vectors.

Dimension: The dimension of U is the number of vectors in a basis of U. It is written as $\dim(U)$.

Finding Dimension and Span: Given a set of vectors $\{\vec{u_m}\}$, we can solve $\begin{bmatrix} \vec{u_1} \cdots \vec{u_m} \mid \vec{0} \end{bmatrix}$ In row echelon form, redundant columns can be eliminated such that the associated matrix in row echelon form is full rank. Thus, $\dim(U)$ is the rank of the matrix, or the number of vectors in the span.

Nullspace $N(A)$

Definition: Let A be an $m \times n$ matrix. The nullspace of A is:

$$N(A) = \left\{ \vec{x} \in \mathbb{R} : A\vec{x} = \vec{0} \right\}$$

In plain English, it is said that the nullspace is the set of vectors which by multiplication of A turn into the zero vector.

Finding the Nullspace: The nullspace can be obtained by solving the matrix $\begin{bmatrix} A \mid \vec{0} \end{bmatrix}$, such that $A\vec{x} = \vec{0}$. The span of the solution to the matrix is the nullspace. The number of vectors in the span is the dimension of the matrix.

Range $R(A)$

Definition: Let A be an $m \times n$ matrix. The range of A is:

$$R(A) = \left\{ \vec{y} \in \mathbb{R} : A\vec{x} = \vec{y} \text{ for some } \vec{x} \in \mathbb{R}^M \right\}.$$

Finding the Range: Given a matrix A, use Gaussian elimination to bring it into row echelon form. The columns in the original matrix who have pivot entries in the REF matrix form the span of the range.

Rank Nullity Theorem

Let U be the row echelon form of A. We can generalize that:

$\dim(R(A))$: Number of columns in U with a leading nonzero pivot element (rank of the matrix).

$\dim(N(A))$: Number of columns in U without a leading nonzero pivot element.

Theorem: The rank-nullity theorem states that $\dim(R(A)) + \dim(N(A)) = n$ for an $n \times m$ matrix.

Implications of the Rank-Nullity Theorem

Theorem: Let $A = LU$ be the LU decomposition of A (if it exists), and let $r = \text{rank}(A)$.

Then, $R(A) = \text{span}\{\vec{l_1}, \cdots, \vec{l_r}\}$ where $\vec{l_n}$ are the first r columns of L.