Course Information

1. Meeting time: 9.00-11.00 and 11.15-1.15.

2. Course Syllabus
   (a) Linear Algebra
      i. Gaussian Elimination.
      ii. Vector Spaces and the Fundamental Subspaces.
      iii. Orthogonality.
      iv. Determinants.
      v. Eigenvalues and Eigenvectors.
      vi. Positive Definite Matrices, Norms and the SVD.
      vii. Linear Programming using the Simplex Method.
   (b) Probability refresher.
References


Section I: Gaussian Elimination

1. Geometry of Linear Equations
2. An example of Gaussian Elimination
3. LU decomposition
4. Gauss-Jordan elimination and Inverses
Introduction

1. Central Problem: \( n \) linear equations in \( n \) unknowns.
2. Methods:
   (a) Elimination to convert system to a system of \( n - 1 \)
       equations in \( n - 1 \) unknowns. In particular, we shall
       discuss issues such as:
       i. Geometry of Linear Equations.
       ii. Factorization.
       iii. Failure of Elimination Procedures.
       iv. Operation counts.
   (b) Method of Determinants (Cramer’s Rule).
   (c) Iterative Methods.

Row Space Geometry

1. We start with an example in two dimensional space.
   \[
   \begin{align*}
   2x - y &= 1 \\
   x + y &= 5
   \end{align*}
   \]

2. In matrix notation, this becomes
   \[
   \begin{pmatrix}
   2 & -1 \\
   1 & 1
   \end{pmatrix}
   \begin{pmatrix}
   x \\
   y
   \end{pmatrix}
   =
   \begin{pmatrix}
   1 \\
   5
   \end{pmatrix}
   .
   \]

3. Such systems can be solved:
   (a) Systematic elimination of variables to solve system.
   (b) Each equation represents a plane in \( n \) dimensional
       space.
   (c) Solution represents intersection of \( n \) planes.
Column Space Geometry

1. \( x \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \).

2. Combination of column vectors that produce vector on right side.

3. Coefficients of column vectors represent solution of system.

Singular Systems

1. Possibilities
   (a) Inconsistent Systems of Equations.
   (b) Infinity of Solutions.

2. If \( n \) planes have no point in common, \( n \) columns lie in the same plane.

3. Row-system has no solution \( \equiv \) Column system has no solution.
A Short Example

1. An inconsistent system.

\[
\begin{align*}
    u + v + w &= 2, \\
    2u + 2w &= 5, \\
    3u + v + 3w &= 6.
\end{align*}
\]

2. Eqn. 1 + Eqn 2 - Eqn. 3 \(\implies\) 0 = 1.

3. A consistent system.

\[
\begin{align*}
    u + v + w &= 2, \\
    2u + 2w &= 5, \\
    3u + v + 3w &= 7.
\end{align*}
\]

4. Eqn. 1 + Eqn 2 - Eqn. 3 \(\implies\) 0 = 0.

A Short Example: II

1. Essentially 3 planes have a whole line in common and we have fewer equations than unknowns \(\implies\) infinity of solutions.

2. Column space picture is as follows:
   (a) All column vectors lie in same plane and are linearly dependent.
   (b) If rhs vector is not in the same plane \(\implies\) inconsistency.
   (c) If rhs vector is in same plane \(\implies\) infinity of solutions.
An Example of Gaussian Elimination

1. Consider the following example of a system of 3 equations in 3 unknowns and the steps that one takes using Gaussian Elimination.

\[
\begin{align*}
2u + v + w &= 5 \\
4u - 6v &= -2 \\
-2u + 7v + 2w &= 9 \\
\rightarrow \\
2u + v + w &= 5 \\
-8v - 2w &= -12 \\
+8v + 3w &= 14 \\
\end{align*}
\]

An Example of Gaussian Elimination

2. \rightarrow

\[
\begin{align*}
2u + v + w &= 5 \\
-8v - 2w &= -12 \\
w &= 2 \\
\end{align*}
\]

3. The forward elimination steps in matrix notation.

\[
\begin{bmatrix}
2 & 1 & 1 & 5 \\
4 & -6 & 0 & -2 \\
-2 & 7 & 2 & 9
\end{bmatrix}
\rightarrow
\begin{bmatrix}
2 & 1 & 1 & 5 \\
0 & -8 & -2 & -12 \\
0 & 0 & 1 & 2
\end{bmatrix}
\]
An Example of Gaussian Elimination

5. The backward substitution phase follows to determine the values of $u$ and $v$. In other words, the solution of $w$ is substituted in the second equation to obtain $v$ and finally $v$ and $w$ are used in the first equation to solve for $u$.

Breakdown of Elimination

1. Forward elimination takes most of computational effort.
2. At each step, zeros are created below the active pivot.
3. Pivots cannot be zero by definition since it is impossible to create zeros below it.
4. Breakdown always occurs if system is singular, usually curable if nonsingular. An example of a singular system is provided below:
   \[ u + v + w = \ldots. \]
5. \[ 2u + 2v + 5w = \ldots. \rightarrow \]
   \[ 4u + 4v + 8w = \ldots. \]
   \[ u + v + w = \ldots. \]
   \[ + 3w = \ldots. \]
   \[ + 4w = \ldots. \]
Complexity of Elimination

1. How many arithmetical operations are required to solve \( n \) equations in \( n \) unknowns using GE?

2. Two types of operations:
   (a) Division by pivot.
   (b) Subtraction of equations.

3. For each equation of length \( k \), \( k^2 - k \) operations to create a column of zeros:
\[
\sum_{k=1}^{n}(k^2 - k) = \frac{n^3 - n}{3}.
\]

4. Back substitution requires \( \sum_{k=1}^{n}k = \frac{n(n+1)}{2} \).

Matrix-Vector Multiplication

1. By Rows: \( Ax = \)
\[
\begin{bmatrix}
1 & 1 & 6 \\
3 & 0 & 3 \\
1 & 1 & 4
\end{bmatrix}
\begin{bmatrix}
2 \\
5 \\
0
\end{bmatrix}
= \begin{bmatrix}
1.2 + 1.5 + 6.0 \\
3.2 + 0.5 + 3.0 \\
1.2 + 1.5 + 4.0
\end{bmatrix}
= \begin{bmatrix}
7 \\
6 \\
7
\end{bmatrix}
\]

2. By Columns: \( Ax = 2 \)
\[
\begin{bmatrix}
1 \\
3 \\
1
\end{bmatrix} + 5 \begin{bmatrix}
1 \\
0 \\
1
\end{bmatrix} + 0 \begin{bmatrix}
6 \\
3 \\
4
\end{bmatrix}
= \begin{bmatrix}
7 \\
6 \\
7
\end{bmatrix}
\]

3. \( \sum_{j=1}^{n} a_{ij} x_j \) is the \( i^{th} \) component of \( Ax \).
Elimination Matrices

1. Given \( A \), how does one create a zero below first pivot?
\[
A = \begin{bmatrix}
2 & 1 & 1 \\
4 & -6 & 0 \\
-2 & 7 & 2
\end{bmatrix}
\]

2. If Row 1 is multiplied by 2 and subtracted from Row 2, we get a zero below the first pivot. This can be achieved by multiplying \( A \) by an elimination matrix, \( E \).
\[
EA = \begin{bmatrix}
1 & 0 & 0 \\
-2 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} \times \begin{bmatrix}
2 & 1 & 1 \\
4 & -6 & 0 \\
-2 & 7 & 2
\end{bmatrix} = \begin{bmatrix}
2 & 1 & 1 \\
0 & -8 & 2 \\
-2 & 7 & 2
\end{bmatrix}
\]

Elimination Matrices: II

1. Note that this works for just 1 column as well.
\[
EA_1 = \begin{bmatrix}
1 & 0 & 0 \\
-2 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} \times \begin{bmatrix}
2 \\
4 \\
-2
\end{bmatrix} \text{ results in } \begin{bmatrix}
2 \\
0 \\
-2
\end{bmatrix}
\]

2. Matrix Multiplication Properties:
   (a) Associative \( \implies (AB)C = A(BC) \).
   (b) Distributive \( \implies A(B + C) = AB + AC \).
   (c) Not always Commutative \( \implies FE \neq EF \).
Triangular Factors and Row Exchange

1. We shall now discuss the appearance of triangular factors during the elimination process.

\[ Ax = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix} = b \]

2. After proceeding with elimination, we are left with the upper triangular system:

\[ Ux = \begin{bmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 5 \\ -12 \\ 2 \end{bmatrix} = c \]

4. The elimination can be represented as product of the following elimination matrices: \( GFE = \)

\[ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \]

5. Reversal of particular elimination step can be achieved quite simply by taking the inverse of this product of the elimination matrices:

\[ \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]
Triangular Factors and Row Exchange

6. We can recover $A$ from $U$ by pre-multiplying it by the inverses of the elimination matrices in reverse order. In fact, the reversed product of the inverses is a lower triangular matrix that forms the $'L'$ of the $LU$ decomposition.

7. $E^{-1}F^{-1}G^{-1}U = A$ where $E^{-1}F^{-1}G^{-1}$

\[
\begin{bmatrix}
1 & 0 & 0 \\
2 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
-1 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & -1 & 1
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 \\
2 & 1 & 0 \\
-1 & -1 & 1
\end{bmatrix} = L
\]

Triangular Factorization: A Recap

1. If no row permutations, $A = LU$.
2. Given $L, U$ solve 2 Triangular systems.
3. $Lc = b$.
4. $Ux = c$.
5. $A$ is factorizable in $\frac{n^3}{3}$ steps.
6. Triangular Systems solvable in $\frac{n^2}{2}$ steps.
An Example of Elimination

1. \( A = \begin{bmatrix}
1 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 2 \\
\end{bmatrix} = LU = \begin{bmatrix}
1 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & -1 & 1 \\
\end{bmatrix} \begin{bmatrix}
1 & -1 & 0 & 1 \\
0 & 1 & -1 & 0 \\
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}\)

An Example of Elimination:II

1. \( Ax = b, \begin{bmatrix}
x_1 & - x_2 \\
- x_1 & +2x_2 & - x_3 \\
& - x_2 & +2x_3 & +x_4 \\
& & - x_3 & +2x_4 \\
\end{bmatrix} = 1 \)

2. \( Lc = b, \begin{bmatrix}
c_1 \\
- c_1 & + c_2 \\
& - c_2 & + c_3 \\
& & - 2 c_3 & + c_4 \\
\end{bmatrix} = 1 \) which results in

\( c = \begin{bmatrix}
1 \\
2 \\
3 \\
4 \\
\end{bmatrix} \).
An Example of Elimination: III

1. \( Ux = c, \begin{bmatrix} x_1 & -x_2 & & & = 1 \\ + x_2 & -x_3 & = 1 \\ + x_3 & -x_4 & = 1 \\ + x_4 & = 1 \end{bmatrix} \) which results in

\[
x = \begin{bmatrix} 10 \\ 9 \\ 7 \\ 4 \end{bmatrix}.
\]

Permutation Matrices

1. Recall that sometimes one is faced with a zero pivot to begin with and the rows have to be permuted to obtain a nonzero pivot.

2. A simple example is the system \( \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \).

3. The first pivot is zero and one can remedy the situation quite simply by exchanging equations 1 and 2. This would result in:

\[
\begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} b_2 \\ b_1 \end{bmatrix}.
\]
Permutation Matrices: II

1. This exchange or permutation of equations can be achieved by the permutation matrix: \[
\begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}.
\]

2. At every stage there may be a permutation required and the resulting permutations will have an impact on the final $L$ and $U$ factors that result. In particular, $P_nP_{n-1}\ldots P_1A = LU$ or just $PA = LU$.

Inverses and Transposes

1. If $Ax = b$ then $A^{-1}b = x$. **Never Calculate the Inverse Directly**

2. $A$ is invertible if there exists a matrix $B$ such that $AB = BA = I$. There is at most one such matrix and is denoted by $A^{-1}$.

3. (a) $A^{-1}A = I$.
   
   (b) $(AB)^{-1} = B^{-1}A^{-1}$.

4. Calculation of $A^{-1}$
   
   (a) Consider equation $AA^{-1} = I$, $A \in \mathbb{R}^{n \times n}$
   
   (b) This can be decomposed into $AA^{-1}_j = I_j$, $j \in 1 \ldots n$.
   
   (c) Each of these equations can be solved using GE to obtain columns of $A^{-1}$ by $LUx_j = e_j$. 
Section I: Review

1. Row Picture: Intersection of Planes.
2. Column Picture: The rhs $b$ is a combination of column vectors.
3. Intersection point of planes = coefficients in combination of columns.
4. Inconsistency: $b$ cannot be formed by combination of columns $\implies$ no solution.
5. Singularity: Planes all have a line in common $\implies$ infinite solutions.
6. If $A$ is nonsingular, then $Ax = b$ has a unique solution, $P$ is a permutation matrix that reorders the rows ensuring no zero pivots.
7. If $A$ is singular, no reordering can produce a full set of pivots.

Section II: The Fundamental Subspace

1. Vector spaces and subspaces
2. The solution of $m$ equations in $n$ unknowns.
3. Linear Independence
4. The four fundamental subspaces
5. Existence and Uniqueness of Inverses
Vector Spaces and Subspaces

1. A real vector space is a set of vectors along with the following rules for vector addition and multiplication by scalars.
   (a) \( x + y = y + x. \)
   (b) \( x + (y + z) = (x + y) + z. \)
   (c) There is a unique "zero vector" such that \( x + 0 = x. \)
   (d) For each \( x, \) there is a unique vector \(-x\) such that \( x + (-x) = 0. \)
   (e) \( 1x = x. \)
   (f) \( (c_1c_2)x = c_1(c_2x). \)
   (g) \( c(x + y) = cx + cy. \)
   (h) \( (c_1 + c_2)x = c_1x + c_2x. \)

Vectors Spaces and Subspaces: II

1. Some examples of vector spaces are:
   (a) \( \mathbb{R}^\infty. \)
   (b) \( \mathbb{R}^{3\times2}. \)

2. A subspace of a vector space is a nonempty subset that satisfies two requirements:
   (a) If we add any \( x \) and \( y \) in the subspace, the sum lies in the subspace.
   (b) If we multiply a vector \( x \) by a scalar \( c, \) the vector \( cx \) lies in the subspace.
Solvability of Systems

1. The system $Ax = b$ is solvable if and only if the vector $b$ can be expressed as a combination of the columns of $A$.

2. Consider the following system:

$$
\begin{bmatrix}
1 & 0 \\
5 & 4 \\
2 & 4 \\
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
\end{bmatrix}
=
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
\end{bmatrix}
$$

3. This can be rewritten as:

$$
\begin{bmatrix}
1 \\
5 \\
2 \\
\end{bmatrix}u
+
\begin{bmatrix}
0 \\
4 \\
4 \\
\end{bmatrix}v
=
\begin{bmatrix}
b_1 \\
b_2 \\
b_3 \\
\end{bmatrix}
$$

Column Spaces

1. The column space is the space containing all possible combinations of columns and is denoted by $\mathcal{R}(A)$.

2. $Ax = b$ can be solved if and only if $b$ lies in $\mathcal{R}(A)$.

3. Column Space of $A$ is a 2 dimensional space in a 3 dimensional space (each vector has 3 components).

4. Remaining dimension is a space perpendicular to the column space: In this case this represents a space of 1 dimension (a line).

5. Any $n \times n$ nonsingular matrix will have $\mathbb{R}^n$ as its column space.

6. If we also include singular and rectangular matrices, the column space lies somewhere between the zero space and the whole space.
Nullspaces

1. The set of solutions to $Ax = 0$ is a vector space called the nullspace: $\mathcal{N}(A)$.

2. Just as the column space is a subspace of $\mathbb{R}^m$ for a matrix of dimension $\mathbb{R}^{m \times n}$, the nullspace of $A$ is a subspace of $\mathbb{R}^n$.

3. The nullspace of $A$ is given by

$$\begin{bmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

4. Solution to this is the zero vector (only).

Nullspaces: II

1. Consider the addition of a column to $A$ and let the new matrix be called $B$.

$$B = \begin{bmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{bmatrix}.$$

2. $\mathcal{N}(B) = cx, x = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, c \in \mathbb{R}$.

3. Later we shall delve deeper into the four fundamental subspaces: Column Space, $(\text{Column Space})^\perp$, Null Space and $(\text{Null Space})^\perp$. 

Solution of \( m \) Equations in \( n \) Unknowns

1. Consider the scalar equation \( ax = b \).
   (a) If \( a \neq 0 \), then for any \( b \), a solution \( x = b/a \) exists and is unique. This is called the nonsingular case.
   (b) If \( a = 0 \) and \( b = 0 \), infinite solutions exist and this is called the underdetermined case.
   (c) If \( a = 0 \) and \( b \neq 0 \), there is no solution and this is called the inconsistent case.
   (d) For square matrices, all these alternatives may occur.

2. To any \( m \) by \( n \) matrix \( A \), there corresponds a permutation matrix \( P \), a lower triangular matrix \( L \) with unit diagonal and an \( m \times n \) upper echelon matrix \( U \), such that \( PA = LU \).

Solution of \( m \) Equations in \( n \) Unknowns

3. An echelon form:

\[
\begin{bmatrix}
* & \ldots & * \\
0 & 0 & * \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

where \( * \) represents a pivot.

4. Consider the system:

\[
Ux = \begin{bmatrix}
1 & 3 & 3 & 2 \\
0 & 0 & 3 & 1 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
w \\
y
\end{bmatrix} = \begin{bmatrix}0 \\
0 \\
0
\end{bmatrix}
\]

5. Variables are classified as:
   (a) Basic: Columns with Pivots (1 and 3).
   (b) Free: Columns without Pivots (2 and 4).
Solution of \( m \) Equations in \( n \) Unknowns

6. Method of finding all solutions to \( Ax = 0 \):
   (a) Given \( Ux = 0 \), identify basic and free variables.
   (b) Set one free variable at 1 and others at zero and calculate values of basic variables.
   (c) Repeat step (b) for every free variable. The combinations of all these solutions result in the nullspace of \( A \).

7. Given an \( m \times n \) system where \( m < n \), then there can be at most \( m \) pivots(basic variables) and at least \( n - m \) free variables.

Rectangular Systems of the form \( Ax = b \)

1. Every solution to \( Ax = b \) is a sum of a particular solution and a homogeneous solution.

2. \( x_{\text{general}} = x_{\text{particular}} + x_{\text{homogeneous}} \).

3. \( x_{\text{particular}} \) is obtained by setting all free variables to zero and solving for the basic variables.

4. Thus the algorithm can be stated as follows:
   (a) Reduce \( Ax = b \) to \( Ux = c \).
   (b) Set all free variables to zero and find the particular solution.
   (c) Set rhs to zero and with one free variable at one, find a solution with all other free variables at zero, this is a homogeneous solution.
Rectangular Systems of the form $Ax = $

5. Elimination clarifies the number of pivots and free variables.
6. If there are $r$ pivots, then there are $n - r$ free variables.
7. The number $r$ is called the rank of the matrix
   (a) If $r = n$, there are no free variables in $x$.
   (b) If $r = m$, there are no zero rows in $U$.

Linear Independence and Basis

1. The Rank of the matrix counts the number of linearly independent rows in the matrix.
2. If $c_1v_1 + c_2v_2 + \ldots + c_kv_k = 0$ only if $c_1 = c_2 = \ldots = c_k = 0$, then the vectors $v_1, v_2, \ldots, v_k$ are linearly independent. Otherwise they are linearly dependent and one of them is a linear combination of the other.
3. If the nullspace contains only the zero vector $\iff$ The columns are linearly independent. An example of such a system is as follows:

$$
\begin{bmatrix}
3 & 4 & 2 \\
0 & 1 & 5 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
c_1 \\
c_2 \\
c_3
\end{bmatrix} = 
\begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}
$$
Linear Independence and Basis

4. A set of $n$ vectors in $\mathbb{R}^m$ must be linearly dependent if $n > m$.
   
   (a) $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \end{bmatrix}$
   
   (b) $A \to U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.
   
   (c) $Uc = 0 \implies c_1 = 1, c_2 = -1, c_3 = 1$.

Linear Independence and Basis

5. If a vector space $V$ consists of all linear combinations of vectors $w_1, w_2, \ldots w_l$, then these vectors span the space. In particular, every vector in $V$ can be expressed as some combination of the $w$'s:
   
   $$v = c_1 w_1 + \ldots + c_l w_l.$$ 

6. Recall that the column space is the space spanned by the columns of the matrix.

7. **Important:** Spanning involves the column space and Independence involves the nullspace. Therefore to check if the columns are linearly independent, solve $Ax = 0$. 
Linear Independence and Basis

8. A basis for a vector space is a set of vectors having two properties:
   (a) The vectors are linearly independent and
   (b) they span the space.

9. A space has an infinite number of bases.

10. Finally a basis is a:
   (a) Maximal independent set: it cannot be made any larger without losing the property of linear independence.
   (b) Minimal spanning set: it cannot be made any smaller and still span the space.

The Four Fundamental Subspaces

1. Subspaces can be defined in two ways:
   (a) A basis of vectors for the subspace may be provided, e.g. the columns of a nonsingular square matrix define the column space.
   (b) A set of constraints that must be satisfied by the vectors in the subspace, e.g. the vectors corresponding to $Ax = 0$. 
The Four Fundamental Subspaces:II

1. The four fundamental subspaces are defined as:
   (a) The column space of $A$ is denoted by $\mathcal{R}(A)$ and lies in $\mathbb{R}^m$.
   (b) The nullspace of $A$ is denoted by $\mathcal{N}(A)$ and lies in $\mathbb{R}^n$.
   (c) The row space of $A$ is denoted by $\mathcal{R}(A^T)$ and lies in $\mathbb{R}^n$.
   (d) The left nullspace of $A$ is denoted by $\mathcal{N}(A^T)$ and lies in $\mathbb{R}^m$.

The Four Fundamental Subspaces:III

1. Consider $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

2. The four fundamental subspaces for the matrix above are defined as:
   (a) $\mathcal{R}(A)$ is the line through $\begin{bmatrix} 1 \\
   0 \end{bmatrix}$.
   (b) $\mathcal{R}(A^T)$ is the line through $\begin{bmatrix} 1 \\
   0 \\
   0 \end{bmatrix}$.
The Four Fundamental Subspaces:IV

(c) $\mathcal{N}(A)$ is the column space of
\[
\begin{bmatrix}
0 & 0 \\
1 & 0 \\
0 & 1 
\end{bmatrix}.
\]

(d) $\mathcal{N}(A^T)$ is the column space of
\[
\begin{bmatrix}
0 \\
1
\end{bmatrix}.
\]

Fund. Subspaces: General matrix

1. Consider $A \in \mathbb{R}^{3 \times 4}$,
\[
A = \begin{bmatrix}
1 & 3 & 3 & 2 \\
2 & 6 & 9 & 5 \\
-1 & -3 & 3 & 0
\end{bmatrix} \rightarrow U = \begin{bmatrix}
1 & 3 & 3 & 2 \\
0 & 0 & 3 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

2. A little examination reveals that Row 3 = 2*Row 2 - 5*Row 1 $\implies$ the rows are linearly dependent.

3. The rowspace of $A$ has dimension $r = 2$. Since the rows of $A$ are linear combinations of the rows of $U$, the rowspace of $A$ is equivalent to the rowspace of $U$ and has the same dimension. $\mathcal{R}(A^T)$ is defined as the column space of
\[
\begin{bmatrix}
1 & 0 \\
3 & 0 \\
3 & 3 \\
2 & 1
\end{bmatrix}
\]
Fund. Subspaces: General matrix-II

1. The nullspace of $A$ is the same as the nullspace of $U$. The basis of the nullspace of $U$ can be constructed by setting one free variable at 1 and the others at 0. Since $Ux = 0$ has $n - r = 2$ free variables, the dimension of the nullspace is 2. $\mathcal{N}(A)$ is defined as the column space of

$$
\begin{bmatrix}
-3 & -1 \\
1 & 0 \\
0 & -\frac{1}{3} \\
0 & 1
\end{bmatrix}
$$

Fund. Subspaces: General matrix-III

1. It can be easily verified that any vector in $\mathcal{N}(A)$ is linearly independent with respect to a vector in $\mathcal{R}(A^T)$.

2. The column space is often called the range space. The column spaces of $A$ and $U$ are clearly different. The $\mathcal{R}(U)$ is defined as the column space of

$$
\begin{bmatrix}
1 & 3 \\
0 & 3 \\
0 & 0
\end{bmatrix}
$$
Fund. Subspaces: General matrix-IV

1. If a set of columns of $U$ is independent, then the corresponding set of columns in $A$ is independent and vice versa. The dimension of the $\mathcal{R}(A)$ equals the rank $r$ which also equals the dimension $\mathcal{R}(A^T)$.

2. The number of independent columns is equal to the number of independent rows. A basis for $\mathcal{R}(A)$ is obtained by forming the $r$ columns of $A$ that correspond to the columns of $U$ containing pivots. This is one of the most important theorems in Linear Algebra and is written as $\textbf{Row rank = Column rank}$.

3. The left nullspace of $A$ or $\mathcal{N}(A^T)$ is defined as the space of $y$'s that correspond to $y^T A = 0^T$ or $A^T y = 0$. The dimension of this nullspace is defined by: $\dim(\mathcal{R}(A)) + \dim(\mathcal{N}(A^T)) = \text{number of columns} = n$.

Fund. Subspaces: General matrix-V

1. We also have $\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A^T)) = \text{number of rows} = m$.

2. Finally, a recap:
   (a) $\mathcal{R}(A) =$ column space of $A$; dimension $r$.
   (b) $\mathcal{N}(A) =$ nullspace of $A$; dimension $n - r$.
   (c) $\mathcal{R}(A^T) =$ row space of $A$; dimension $r$.
   (d) $\mathcal{R}(A) =$ left nullspace of $A$; dimension $m - r$. 
Existence and Uniqueness of Inverses

1. Having just discussed rank, we can now make some statements about the existence and uniqueness of inverses.

2. **An inverse exists only when the rank is large as possible;**
   Rank \( r \) satisfies \( r \leq m \) and \( r \leq n \).

3. **Existence:** The system \( Ax = b \) has at least one solution \( x \) for every \( b \iff \) columns span \( \mathbb{R}^m \); then \( r = m \). In this case, there exists a right inverse \( C \in \mathbb{R}^{n \times m} \) such that \( AC = I_m, m \leq n \). The left inverse \( B = (A^T A)^{-1} A^T \).

Existence and Uniqueness of Inverses:

1. **Uniqueness:** The system \( Ax = b \) has at most one solution \( x \) for every \( b \iff \) the columns are linearly independent; then \( r = n \). In this case, there exists a left inverse \( B \in \mathbb{R}^{m \times n} \) such that \( BA = I_m, m \geq n \). The right inverse \( C = A^T (A A^T)^{-1} \).

2. For a rectangular matrix, we cannot have both existence and uniqueness.

3. For a square matrix, existence \( \iff \) uniqueness.

4. For a square matrix of dimension \( n \), following conditions are necessary and sufficient for the matrix to be nonsingular:
Existence and Uniqueness of Inverses:

1. The columns span \( \mathbb{R}^n \) so \( Ax = b \) has at least one solution for every \( b \).
2. The columns are linearly independent, so \( Ax = 0 \) has only one solution \( x = 0 \).
3. The rows of \( A \) span \( \mathbb{R}^n \).
4. The rows are linearly independent.
5. Elimination can be completed: \( PA = LDU \), with all \( d_i \neq 0 \).
6. There exists a matrix \( A^{-1} \) such that \( AA^{-1} = A^{-1}A = I \).
7. \( Det(A) \neq 0 \).
8. Zero is not an eigenvalue of \( A \).
9. \( A^TA \) is positive definite.

Section II: Review

1. \( Ax = b \) is solvable only if \( b \in \mathcal{R}(A) \).
2. For any \( m \times n \) rectangular matrix, \( PA = LU \), where \( U \) is upper echelon.
3. \( r \) pivots \( \implies \) \( r \) basic variables and \( n - r \) non-basic variables.
4. Rank of the matrix is \( r \).
5. If \( r = n \), \( \implies \) no free variables and nullspace is empty.
6. If \( r = m \) \( \implies \) no zero rows in \( U \) and system can be solved for any \( b \).
Section II: Review: II

1. $\mathcal{N}(A) \cup \mathcal{R}(A^T) = \mathbb{R}^n$.
2. $\mathcal{N}(A^T) \cup \mathcal{R}(A) = \mathbb{R}^m$.
3. Fundamental Theorem: **Row Rank = Column Rank**.
4. Fundamental Theorem-II
   (a) $\mathcal{R}(A) \in \mathbb{R}^r$
   (b) $\mathcal{N}(A) \in \mathbb{R}^{n-r}$
   (c) $\mathcal{R}(A^T) \in \mathbb{R}^r$
   (d) $\mathcal{N}(A^T) \in \mathbb{R}^{m-r}$

Section III: Orthogonality

1. Perpendicular Vectors and Orthogonal Subspaces
2. Inner Products and Projections onto lines
3. Projections and Least Squares Approximations
4. Orthogonal Bases and the Gram Schmidt Orthogonalization
Perp. Vectors and Orthog. Subspaces

1. The length $\|x\|$ of a vector $x$ in $\mathbb{R}^n$ is the positive square root of $\|x\|^2 = x_1^2 + \ldots + x_n^2 = x^T x$.
2. $x$ and $y$ are orthogonal if $x_1y_1 + \ldots + x_ny_n = 0$.
3. If the nonzero vectors $v_1, \ldots, v_k$ are mutually orthogonal then they are linearly independent.
4. An important example of a set of mutually orthogonal set of vectors is the set of coordinate unit vectors $e_1, \ldots, e_n$. This system can be rotated to produce a new orthonormal basis.

Orthogonal Subspaces

1. Two subspaces $V$ and $W$ are orthogonal if every vector $v$ in $V$ is orthogonal to every vector $w$ in $W$.
2. Our interest is in orthogonal subspaces:
   (a) The row space is orthogonal to the nullspace in $\mathbb{R}^n$.
   (b) The column space is orthogonal to the left nullspace in $\mathbb{R}^m$.
   (c) Idea of proof:
      i. Suppose $x$ is in the nullspace and $v$ is in the row space.
      ii. $Ax = 0$ and $v = A^T z$ for some vector $z$.
      iii. $v^T x = (A^T z)^T x = z^T A x = z^T 0 = 0$.
3. Given a subspace $V$ of $\mathbb{R}^n$, the space of all vectors orthogonal to $V$ is called the orthogonal complement of $V$ and denoted by $V^\perp$. 
Orthogonal Subspaces: II

1. **Fundamental Theorem of Linear Algebra**: II
   (a) The nullspace is the orthogonal complement of the row space in \( \mathbb{R}^n \) or \( \mathcal{R}(A^T) = (\mathcal{N}(A))^\perp \).
   (b) The left nullspace is the orthogonal complement of the column space in \( \mathbb{R}^n \) or \( \mathcal{N}(A^T) = (\mathcal{R}(A))^\perp \).

2. \( Ax = b \) is solvable \( \iff \) \( b^T y = 0 \) whenever \( A^T y = 0 \).

3. Effect of the transformation of \( A \):
   (a) \( x \in \mathbb{R}^n, x = x_{\text{rowspace}} + x_{\text{nullspace}} \).
   (b) Transformation: \( Ax = Ax_r + Ax_n \)
   (c) \( Ax_n = 0 \), Nullspace component \( \rightarrow \) 0.
   (d) \( Ax_r = Ax \), Row space component \( \rightarrow \) column space.
   (e) Every matrix transforms its row space(\( \mathbb{R}^n \)) to its column space(\( \mathbb{R}^m \)).

---

Projections onto subspaces

1. Given a point \( b \) in \( \mathbb{R}^n \), we want to find the point closest to a line (or possibly a subspace) \( a \), say \( p \). The point \( p \) is called the projection of \( b \) onto \( a \). The line connecting \( b \) to \( a \) is perpendicular to \( a \).

2. Applications: Least Squares Solution to an Overdetermined System (more equations than unknowns and possibly inconsistent).

3. Angle between \( b \) and \( a \) is given by \( \cos \theta = \frac{a^T b}{||a|| ||b||} \).

4. Let \( p = \bar{x}a \), \( \bar{x} = \frac{a^T b}{a^T a} \).

5. \( p = \frac{a^T b}{a^T a} a \).

6. **Schwartz Inequality**: \( |a^T b| \leq ||a|| ||b|| \), where equality holds if and only if \( b \) is a multiple of \( a \).
Projection Matrices

1. Projection onto a line is carried out by a Projection Matrix.

2. It is a matrix that multiplies $b$ and produces $p$.

3. $P = \frac{aa^T}{a^T a}$.

4. A projection onto line through $a = (1,1,1)$ and is given by

$$
\begin{bmatrix}
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3}
\end{bmatrix}
$$

Projection Matrices: II

1. Properties of $P$:
   (a) $P$ is a symmetric matrix.
   (b) $P^2 = P$.

2. This example is also useful in understanding the fundamental subspaces:
   (a) The rank of $P$ is $r = 1$.
   (b) The column space of $P$ is identical to the row space of $P$.
   (c) The column space consists of the line through $a = (1,1,1)$.
   (d) The nullspace of $P$ consists of a plane perpendicular to $a$. 
Projections and Least Squares

1. We shall start with an overdetermined example:

\[
\begin{align*}
2x &= b_1 \\
3x &= b_2 \\
4x &= b_3
\end{align*}
\] (1) (2) (3)

2. The solution \( x \) will exist only if \( b \) is on the same line as the vector as \[
\begin{bmatrix}
2 \\
3 \\
4
\end{bmatrix}
\] otherwise the system is inconsistent.

3. Inconsistent equations arise frequently in practice and one needs to determine the \( x \) that minimizes the average error in the \( m \) equations.

Projections and Least Squares: II

1. The least squares problem can be represented as:

\[
\min E^2 = \|ax - b\|^2 = (a_1x - b_1)^2 + \ldots + (a_mx - b_m)^2.
\]

2. The least squares solution to the above problem in a single unknown is \( x = \frac{a^Tb}{a^Ta} \).
Multivariate Least Squares Problems

1. We shall now consider the more difficult problem of projecting \( b \) onto a subspace.

2. This can again be looked at as a problem of minimizing \( E = ||Ax - b|| \).

3. Geometrically, this can be looked at as measuring the distance from \( b \) to \( Ax \) in the column space.

4. Moreover, the vector \( b - Ax \) is perpendicular to the column space of \( A \) or it lies in the left null space of \( A \).

5. This can be represented as

\[
A^T(b - Ax) = 0. \tag{4}
\]

\[
A^T Ax = A^T b. \tag{5}
\]

Multivariate Least Squares Problems:

1. This is the same as the solution to the minimization problem \( E^2 = (Ax - b)^T(Ax - b) \).

2. The least squares equations \( A^T Ax = A^T b \) are known in statistics as the **Normal Equations**.

3. The least squares solution to an inconsistent system \( Ax = b \) of \( m \) equations in \( n \) unknown satisfies:

\[
A^T Ax = A^T b.
\]

\[
x = (A^T A)^{-1} A^T b.
\]

\[
p = Ax = A(A^T A)^{-1} A^T b.
\]
Normal Equations

1. Let us examine some special cases of the normal equations:
   (a) Suppose \( b \) lies in the column space of \( A \).
   Then we have \( p = Ax = A(A^T A)^{-1} A^T b = A(A^T A)^{-1} A^T A x = Ax = b \).
   (b) Suppose \( b \) is perpendicular to the column space.
   \( A^T b = 0 \),
   (c) When \( A \) is square and invertible, the column space is the whole space and the minimal error is zero.
   \( p = Ax = A(A^T A)^{-1} A^T b = A A^{-1}(A^T)^{-1} A^T b = b \).

2. Properties of \( A^T A \).
   (a) \( A^T A \) has the same nullspace as \( A \).
   (b) If \( A \) has linearly independent columns, then \( A^T A \) is square, symmetric and invertible.

Normal Equations: II

1. Properties of \( A(A^T A)^{-1} A^T \).
   (a) \( P^2 = P \)
   (b) \( P^T = P \)

2. Finally, if \( A \) is actually invertible, the column space spans the whole space and
   \( P = A(A^T A)^{-1} A^T = A A^{-1}(A^T)^{-1} A^T = I \).
Orthogonal Bases and Matrices

1. The vectors $q_i$ and $q_j$ are orthonormal if

$$q_i^T q_j = 0, \ i \neq j,$$

$$q_i^T q_j = 1, \ i = j.$$

2. The standard basis are defined below: $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$, \ldots, $e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

Orthogonal Bases and Matrices: II

1. This is not the only orthonormal basis and can be rotated in any direction while maintaining perpendicularity.

2. An orthogonal matrix is simply a square matrix with orthonormal columns, in particular:

$$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$Q^T Q = I \quad \implies Q^T = Q^{-1}.$$
Orthogonal Bases and Matrices: III

1. Some examples of Orthogonal Matrices:

   (a) Rotation matrix: \[
   \begin{bmatrix}
   \cos \theta & -\sin \theta \\
   \sin \theta & \cos \theta
   \end{bmatrix}.
   \]

   (b) Permutation matrix: \[
   \begin{bmatrix}
   0 & 1 & 0 \\
   0 & 0 & 1 \\
   1 & 0 & 0
   \end{bmatrix}.
   \]

2. Some other properties of orthogonal matrices:

   (a) Multiplication by an orthogonal matrix \(Q\) preserves lengths: \[||Qx||^2 = x^TQ^TQx = x^Tx = ||x||^2 \forall x.\]

   (b) Multiplication by an orthogonal matrix \(Q\) preserves inner products: \((Qx)^T(Qy) = x^Ty, \forall x.\)

Orthogonalization

1. Given a vector \(b\), we want to determine the components in the \(n\) orthogonal directions.

2. \(b\) can be written as \(x_1q_1 + x_2q_2 + \ldots + x_nq_n.\)

3. \(x_1\) can be computed by multiplying both sides by \(q_1\)
   \[\implies q_1^Tb = x_1q_1^Tq_1 = x_1.\]

4. Any vector \(b\) is equal to \((q_1^Tb)q_1 + (q_2^Tb)q_2 + \ldots + (q_n^Tb)q_n.\)

5. This can also be written as
   \[
   x = Q^Tb = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} b \end{bmatrix} = \begin{bmatrix} q_1^Tb \\ q_2^Tb \\ \vdots \\ q_n^Tb \end{bmatrix}. \quad (6)
   \]
Rect. Matrices with OrthonormalCols:

1. Consider the problem of solving $Qx = b$, with $m > n$ - an overdetermined system.
2. We shall solve this by using least squares.
3. If $Q$ has orthonormal columns, then the least squares problem:

$$ Q\bar{x} = b. \quad (7) $$
$$ QTQ\bar{x} = QTb. \quad (8) $$
$$ \bar{x} = QTb. \quad (9) $$
$$ p = Q\bar{x}. \quad (10) $$
$$ p = QQ^Tb. \quad (11) $$

Rect. Matrices with OrthonormalCols:

1. The last formula provides the projection of $b$ onto the column space of $Q$. Remember that the original projection formula reduces to this when $Q^TQ = I$. Moreover, $QQ^T$ is a projection matrix.
Gram-Schmidt Process

1. Given 3 orthogonal vectors, $a$, $b$ and $c$, the projection of $v$ on these three vectors is given quite simply:
   \[ v = (a^T v) a + (b^T v) b + (c^T v) c. \]
2. If we can only assume independence for $a$, $b$ and $c$, can we generate 3 orthogonal vectors?
3. One method to achieve this is through the Gram-Schmidt process:
   (a) \[ q_1 = \frac{a}{\|a\|}. \]
   (b) \[ b' = b - (q_1^T b) q_1. \]
   (c) \[ q_2 = \frac{b'}{\|b'\|}. \]
   (d) \[ c' = c - (q_1^T c) q_1 - (q_2^T c) q_2. \]
   (e) \[ q_3 = \frac{c'}{\|c'\|}. \]

Gram-Schmidt Process: II

1. The Gram-Schmidt Process starts with a set of independent vectors and ends with a set of orthonormal vectors.
QR Factorization

1. We shall now examine the relationship between $a$, $b$ and $c$ and the orthonormal vectors $q_1$, $q_2$ and $q_3$.

2. The Gram-Schmidt process can be represented as

$$
\begin{bmatrix}
a & b & c \\
\end{bmatrix} =
\begin{bmatrix}
q_1 & q_2 & q_3 \\
\end{bmatrix}
\begin{bmatrix}
q_1^T a & q_1^T b & q_1^T c \\
q_2^T b & q_2^T c \\
q_3^T c \\
\end{bmatrix}
$$

3. The matrix of orthonormal vectors is called $Q$ and the upper triangular matrix is called $R$.

QR Factorization: II

1. We shall now provide an example of $QR$ factorization:

$$
A =
\begin{bmatrix}
1 & 1 & 2 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
\end{bmatrix} =
\begin{bmatrix}
1/\sqrt{2} & 1/\sqrt{2} & 0 \\
0 & 0 & 1 \\
1/\sqrt{2} & -1/\sqrt{2} & 0 \\
\end{bmatrix}
\begin{bmatrix}
\sqrt{2} & 1/\sqrt{2} & \sqrt{2} \\
1/\sqrt{2} & \sqrt{2} \\
1 & 1 \\
\end{bmatrix} = QR.
$$

2. Every $m$ by $n$ matrix $A$ with linearly independent columns can be factored into $A = QR$. The columns of $Q$ are orthonormal and $R$ is upper triangular and invertible. When $m = n$ and all matrices are square, $Q$ becomes an orthogonal matrix.
QR Factorization: III

1. Why even orthogonalize? It provides a simplification of the least squares problem:
   \[ A^T A = R^T Q^T Q R = R^T R. \]

2. Thus the normal equations reduce to
   \[
   A^T A \bar{x} = A^T b, \\
   R^T R \bar{x} = R^T Q^T b, \\
   R \bar{x} = Q^T b. \tag{12}
   \]

3. Why is this much better? Since \( R \) is upper triangular, this can be solved quite simply by back-substitution. The major cost lies in the Gram-Schmidt process that creates the \( QR \) factors: it requires \( mn^2 \) operations.

Section III: Review

1. Recall that \( \mathcal{R}(A^T) \perp \mathcal{N}(A) \) and \( \mathcal{R}(A) \perp \mathcal{N}(A^T) \).
2. Every matrix transforms its row space into its column space.
3. Projection of a vector \( b \) onto \( a \) is given by
   \[ p = a(a^T b)/(a^T a). \]
4. The projection matrix \( P = \frac{aa^T}{a^Ta} \) multiplies \( b \) to produce \( p \).
   Moreover \( P = P^T = P^2 \).
5. If \( b \notin \mathcal{R}(A^T) \implies \) the system is inconsistent and GE fails. Least squares (LS) solution chooses the \( x \) that minimizes the error in the \( m \) equations.
6. \( \bar{x} = \frac{a^T b}{a^Ta} \).
7. \( P = A(A^T A)^{-1} A^T \) and projects any vector \( b \) onto the column space of \( A \).
Section III : Review: II

1. $Q$ is orthonormal $\iff Q^T Q = I, Q^T = Q^{-1}$.

2. Multiplication by an orthogonal matrix $Q$ preserves length or $\|Qx\| = \|x\|$.

3. Gram-Schmidt process starts with three vectors $a_1, a_2, \ldots, a_n$ and ends up with $q_1, q_2, \ldots, q_n$ by the factorization $A = QR$, $Q$ orthonormal and $R$ nonsingular and upper triangular.

4. Why?? since it simplifies the least squares process.

Section IV : Determinants

1. Properties of Determinants

2. Some applications of Determinants
Properties of Determinants

1. If $det(A)$ is zero, then $A$ is singular, otherwise $A$ is invertible.

2. The solution to $det(A - \lambda I) = 0$, a $n^{th}$ degree polynomial, provides the $n$ eigenvalues for $A \in \mathbb{R}^{n \times n}$.

3. The determinant of $A$ equals the volume of a parallelepiped $P$ in $n$-space, assuming the edges come from the rows of $A$.

4. The determinant $= \pm$ (product of the pivots).

Properties of Determinants:II

1. $det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

2. The determinant depends linearly on the first row:

$$det \begin{bmatrix} a + a' & b + b' \\ c & d \end{bmatrix} = det \begin{bmatrix} a & b \\ c & d \end{bmatrix} + det \begin{bmatrix} a' & b' \\ c & d \end{bmatrix}.$$ 

$$det \begin{bmatrix} ta & tb \\ c & d \end{bmatrix} = t det \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$
Properties of Determinants: III

1. The determinant changes sign when two rows are exchanged:

\[ \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = -\det \begin{bmatrix} c & d \\ a & b \end{bmatrix}. \]

2. The determinant of the identity matrix is 1.

\[ \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 1. \]

Properties of Determinants: IV

1. If rows of \( A \) are equal, then \( \det(A) = 0 \).

\[ \det \begin{bmatrix} a & b \\ a & b \end{bmatrix} = 0. \]

2. If \( A \) is singular, then \( \det(A) = 0 \) otherwise \( \det(A) \neq 0 \).

3. \( \det(AB) = \det(A) \cdot \det(B) \), \( A, B \) square matrices.

\[ \implies \det(A^{-1}) = \frac{1}{\det(A)}. \]

4. \( \det(A^T) = \det(A) \):

   (a) \( PA = LDU \).

   (b) \( A^T P^T = U^T D^T L^T \).

   (c) \( \det(A^T) \det(P^T) = \det(U^T) \det(D^T) \det(L^T) \).

   (d) But \( L, U \) are triangular and result follows.
Cofactors

1. A determinant of order $n$ can be split into $n$ smaller determinants or minors of order $n - 1$. In particular we have
\[
\det \begin{bmatrix}
    a_{11} & a_{12} & a_{13} \\
    a_{21} & a_{22} & a_{23} \\
    a_{31} & a_{32} & a_{33}
\end{bmatrix}
= a_{11} \det \begin{bmatrix}
    a_{22} & a_{23} \\
    a_{32} & a_{33}
\end{bmatrix}
- a_{12} \det \begin{bmatrix}
    a_{21} & a_{23} \\
    a_{31} & a_{33}
\end{bmatrix}
+ a_{13} \det \begin{bmatrix}
    a_{21} & a_{22} \\
    a_{31} & a_{32}
\end{bmatrix}.
\]

2. Submatrix $M_{1j}$ is formed by removing row 1 and column $j$.

3. Cofactor $A_{1j}$ is defined as
\[
A_{1j} = (-1)^{i+j} \det(M_{1j}).
\]

Cofactors: II

1. Finally we have $\det(A) =$
\[
a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n}.
\]

2. Note that we could also decompose using the $i^{th}$ row
\[
\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}.
\]
Some Applications of Determinants

1. Computation of $A^{-1}$.

2. Recall that $\det(A) = a_{i1}A_{i1} + \ldots + a_{in}A_{in}$.

3. Consider the following matrix product:

$$
\begin{bmatrix}
  a_{11} & a_{12} & \ldots & a_{1n} \\
  a_{21} & a_{22} & \ldots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{n2} & \ldots & a_{nn}
\end{bmatrix}
\begin{bmatrix}
  A_{11} & A_{21} & \ldots & A_{n1} \\
  A_{12} & A_{22} & \ldots & A_{n2} \\
  \vdots & \vdots & \ddots & \vdots \\
  A_{1n} & A_{1n} & \ldots & A_{nn}
\end{bmatrix}
= 
\begin{bmatrix}
  \det(A) & 0 & \ldots & 0 \\
  0 & \det(A) & \ldots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \ldots & \det(A)
\end{bmatrix}.
$$

Some Applications of Determinants: I

1. The diagonal entries follow from the formula above.

2. The off-diagonal entries are equivalent to the following

   $$a_{11}A_{21} + \ldots + a_{1n}A_{2n} = 0.$$ 

3. But this is equivalent to the determinant of a matrix with 2 identical rows.
Some Applications of Determinants: I

1. Therefore we have

\[(A)(A_{cof}) = (\det A)I \implies A^{-1} = \frac{A_{cof}}{\det(A)}\]

if \(\det(A) = 0\), \(A^{-1}\) does not exist.

2. The solution of \(Ax = b\).
   
   (a) \(x = A^{-1}b = \frac{A_{cof}b}{\det(A)}\).

   (b) This can be represented using Cramer’s Rule:

   \[
x_j = \frac{\det B_j}{\det(A)}, B_j = \begin{bmatrix} a_{11} & a_{12} & b_1 & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & b_n & a_{nn} \end{bmatrix}.
   \]

   where vector \(b\) replaces the \(j^{th}\) column of \(A\) in \(B_j\).

Section V: Eigenvalues and Eigenvectors

1. Introduction
2. The diagonal form of a matrix
3. Powers \(A^k\) and Products
4. Similarity Transformations
5. The Jordan Form
An Introduction

1. We shall be considering square matrices unless mentioned otherwise.

2. Why are eigenvalues useful? - shall use an example of differential equations to motivate the definition of eigenvalues.

\[ \frac{dv}{dt} = 4v - 5w, v = 8, t = 0. \]
\[ \frac{dw}{dt} = 2v - 3w, w = 5, t = 0. \]

(13)

An Introduction: II

1. Consider the example of two coupled differential equations:

\[ \frac{dv}{dt} = 4v - 5w, v = 8, t = 0. \]
\[ \frac{dw}{dt} = 2v - 3w, w = 5, t = 0. \]

2. This is an initial value problem and can be stated as follows:

\[ \frac{du}{dt} = Au, u(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}, u(t=0) = \begin{bmatrix} 8 \\ 5 \end{bmatrix}. \]

(14)

3. For the case, where \( A \) is a scalar \( a \), the solution to this is given by \( u(t) = e^{at} u_0. \)
An Introduction: IV

1. Note that for large $t$ the equation is:
   (a) Stable if $a < 0$.
   (b) Neutrally stable if $a = 0$.
   (c) Unstable if $a > 0$.
   (d) We shall not discuss the possibility of the complex case.

An Introduction: V

1. The solutions to the scalar case give us some indication of a solution that could be tried directly. Let us try solutions that have the same exponential dependence on $t$.

   $$v(t) = e^{\lambda t} y.$$  
   $$w(t) = e^{\lambda t} z.$$  
   $$\lambda e^{\lambda t} y = 4 e^{\lambda t} y - 5 e^{\lambda t} z$$  
   $$\lambda e^{\lambda t} z = 2 e^{\lambda t} y - 3 e^{\lambda t} z$$  
   $$4y - 5z = \lambda y$$  
   $$2y - 3z = \lambda z$$
An Introduction: IV

1. \( Ax = \lambda x, A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}. \)

2. This equation \( Ax = \lambda x \) is the fundamental equation of the chapter. The number \( \lambda \) is called the eigenvalue of the matrix and \( x \) is the associated eigenvector. This chapter shall discuss how one goes about determining the eigenvalues and eigenvectors and where one uses them.

Solutions to \( Ax = \lambda x \)

1. The equation \( Ax = \lambda x \) can be written as \( (A - \lambda I)x = 0. \)

2. Such a formulation implies the following:
   (a) The vector \( x \) is in the nullspace of \( (A - \lambda I) \).
   (b) The number \( \lambda \) is chosen so that \( (A - \lambda I) \) has a nullspace.

3. It is obvious that the a matrix always has a nullspace- it could however only contain the zero vector. We are however interested in those values of \( \lambda \) that produce nonzero eigenvectors \( x \). Therefore, \( (A - \lambda I) \) must be singular.

4. \( (A - \lambda I) \) is singular \( \iff \) \( det(A - \lambda I) = 0 \) (characteristic equation or polynomial).
Solutions to $Ax = \lambda x$: II

1. We shall conclude with our earlier example:

$$(A - \lambda I) = \begin{bmatrix} (4 - \lambda) & -5 \\ 2 & (-3 - \lambda) \end{bmatrix}.$$

$$det(A - \lambda I) = 0 \implies \lambda^2 - \lambda - 2 = 0. \quad (15)$$

2. This is called the characteristic polynomial and has solutions $\lambda = -1$ or $\lambda = 2.$

Solutions to $Ax = \lambda x$: III

1. We can now solve for $Ax = \lambda x$ using each value of $\lambda$ resulting in the eigenvectors:

$$\lambda_1 = -1; x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$\lambda_2 = 2; x_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$
Solutions to $A x = \lambda x$: IV

1. To recap:
   (a) Compute the determinant of $(A - \lambda I)$ as a characteristic polynomial in $\lambda$.
   (b) Find the roots of the polynomial: these are the eigenvalues (may be a non-trivial exercise).
   (c) For each eigenvalue solve the equation $A x = \lambda x$. These are the eigenvectors.

2. The solution to the original pair of DEs is given by:

$$u(t) = 3e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + e^{2t} \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$  

A Physical Understanding

1. Eigenvalues and eigenvectors appear naturally when talking about the system $u = e^{\lambda t} x$, in which the eigenvalue gives the rate of growth/decay and the eigenvector changes at this rate.

2. Not all vectors $x$ are eigenvectors (else we would be jobless) unless $A = I$. Eigenvectors are the "normal modes" of the system and act independently. The superposition of these normal modes allows us to construct a solution.

3. The suggested text has dealt with several examples such as difference equations, Fibonacci numbers, Markov processes and differential equations. In the interest of time, we shall not discuss these any further. Before proceeding ahead we shall discuss Diagonalization.
A Physical Understanding: II

1. Finally, two useful properties of eigenvalues are:

\[
\text{Trace}(A) = \lambda_1 + \lambda_2 + \ldots \lambda_n = \\
a_{11} + a_{22} + \ldots + a_{nn}.
\]

\[
\lambda_1 \lambda_2 \ldots \lambda_n = \text{det}(A).
\]

The Diagonal Form of a Matrix

1. Suppose the \( n \) by \( n \) matrix \( A \) has \( n \) linearly independent eigenvectors. Then if these vectors are chosen to be the columns of a matrix \( S \), it follows that \( S^{-1}AS \) is a diagonal matrix \( \Lambda \) with the eigenvalues along its diagonal.

\[
S^{-1}AS = \\
\begin{bmatrix}
\lambda_1 \\
& \lambda_2 \\
& & \ddots \\
& & & \lambda_n
\end{bmatrix}
\]

\[
AS = S\Lambda.
\]

\[
\Lambda = S^{-1}AS.
\]

\[
A = SAS^{-1}.
\]
The Diagonal Form of a Matrix: II

1. If $A$ has no repeated eigenvalues then the $n$ eigenvectors are independent and $A$ can be diagonalized.

2. $S$ is not unique since a multiple of an eigenvector could very well be used.

3. Unless $S$ holds the eigenvectors of $A$ as its columns, the relationship $AS = SA\Lambda$ will not hold.

4. Not all matrices have $n$ linearly independent eigenvectors and therefore not all matrices are diagonalizable. For example $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has repeated eigenvalues equal to 0 (algebraic multiplicity = 2) while its geometric multiplicity is 1 (since it has only one independent eigenvector, $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$).

The Diagonal Form of a Matrix: III

1. The failure of diagonalization is not because the eigenvalues are zero. A matrix could have nonzero eigenvalues and yet not be diagonalizable: $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$.

The problem is the shortness of eigenvectors for constructing $S$.

2. Diagonalizability may fail if there are repeated eigenvalues but not always e.g. the Identity matrix has repeated eigenvalues equal to 1 but there is no shortage of eigenvectors.

3. Finally if the eigenvectors $x_1, x_2, \ldots x_n$ correspond to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then these eigenvectors are linearly independent.
Powers and Products: $A^k$ and $AB$

1. Given the eigenvalues of $A$, the eigenvalues of $A^2$ are merely $\lambda^2$. This can be seen by $A^2 x = A \lambda x = \lambda^2 x$. Moreover the eigenvectors are the same. Note this can be shown by diagonalization as well: $S^{-1} A^2 S = S^{-1} A S S^{-1} A S = \Lambda \Lambda = \Lambda^2$.

2. This can be extended to the $k^{th}$ power and each eigenvector of $A$ is an eigenvector of $A^k$ and if $S$ diagonalizes $A$ then it diagonalizes $A^k$: $\Lambda^k = S^{-1} A^k S$.

3. If $A$ and $B$ are diagonalizable, they share the same eigenvector matrix $S$ if and only if $AB = BA$.

Symmetric Matrices

1. A symmetric matrix has real eigenvalues.

2. The eigenvectors of a symmetric matrix can be chosen to be orthonormal.

3. A real symmetric matrix can be factored into $A = Q \Lambda Q^T$ with the orthonormal vectors in $Q$ and the eigenvalues in $\Lambda$. In particular if the eigenvectors correspond to distinct eigenvalues, then they are orthogonal. This is called the spectral theorem.
Symmetric Matrices

1. We shall provide a proof for this theorem:

   \[ Ax = \lambda_1 x, Ay = \lambda_2 y, A = A^T. \]

   \[ (\lambda_1 x)^T y = (Ax)^T y = x^T A^T y = x^T Ay = \lambda_2 x^T y. \]

   But \( \lambda_1 \neq \lambda_2 \implies x^T y = 0. \)

Similarity Transformations

1. \( A \) and \( S^{-1}AS \) are said to be \textit{similar}.

2. Moving between one and the other is called a similarity transformation.

3. If \( B = S^{-1}AS \), then \( A \) and \( B \) have the same eigenvalues. An eigenvector \( x \) of \( A \) corresponds to an eigenvector \( S^{-1}x \) of \( B \):

   \[ Ax = \lambda x \implies SBS^{-1}x = \lambda x \implies B(S^{-1}x) = \lambda(S^{-1}x). \]

   This can also be confirmed by showing that \( det(A - \lambda I) = det(B - \lambda I) \).
Similarity Transformations

1. **Schur’s lemma**: Given a square matrix $A$, there is an orthogonal matrix $Q$ such that $Q^{-1}AQ = T$ where $T$ is upper triangular. The eigenvalues of $A$ must be shared by the similar matrix $T$ and appear along its main diagonal.

2. If $A$ is symmetric then so is $Q^{-1}AQ$:
   \[(Q^{-1}AQ)^T = QTAT(Q^{-1})^T = Q^{-1}AQ.\]

3. If a symmetric matrix is triangular then it must be diagonal.

4. **Spectral theorem** Every real symmetric matrix can be diagonalized by an orthogonal matrix $\implies Q^{-1}AQ = \Lambda$.

The Jordan Form

1. In the previous section, we made $A$ triangular through a similarity transformation $Q^{-1}AQ$, with the restriction being $Q$ has to be orthogonal.

2. In this section, we shall remove the restriction on $Q$ and try and make $A$ as "diagonal" as possible. The result is the **Jordan Form**.

3. $J = S^{-1}AS = \Lambda$, if $A$ has a full set of independent eigenvectors with $\Lambda$ being a diagonal matrix of eigenvalues.
The Jordan Form: II

1. If $A$ has $s$ independent eigenvectors, it is similar to a matrix with $s$ blocks:

$$J = M^{-1}AM = \begin{bmatrix}
J_1 & & \\
& J_2 & \\
& & \ddots \\
& & & J_n
\end{bmatrix}.$$

$$J_i = \begin{bmatrix}
\lambda_i & 1 \\
& \ddots & 1 \\
& & \lambda_i
\end{bmatrix}.$$

The Jordan Form: III

1. $J_i$ is called a Jordan block and is a triangular matrix with only a single eigenvalue $\lambda_i$ and only one eigenvector. When the block has order $m$, the eigenvalue is repeated $m$ times and has $(m - 1)$ 1’s above the diagonal.

2. if $A = M^{-1}JM$, $A^k = M^{-1}J^kM$. 
Section V: Review

1. The vector $x$ is in the nullspace of $A - \lambda I$.
2. $\lambda$ is chose so that $A - \lambda I$ has a nullspace or that it is singular.
3. Characteristic polynomial $= \det(A - \lambda I) = 0$. Roots are the eigenvalues and the solutions of $Ax = \lambda x$ are eigenvectors for each eigenvalue.
4. If $A$ has $n$ linearly independent eigenvectors, then $S^{-1}AS = \Lambda$. Nonsingularity of $S$ follows from LI of its eigenvectors. Note that $S$ is not unique.
5. Diagonalizability is concerned with eigenvectors.
6. Nonsingularity is concerned with eigenvalues.
7. $\Lambda^k = S^{-1}A^kA \implies \lambda(A^k) = \Lambda^k$.

Section V: Review II

1. If $A$ is real and symmetric then $S$ can be chosen to be orthonormal.
2. $A$ and $M^{-1}AM$ are said to be similar and transformations between them are called similarity transformations. Moreover, they share the same eigenvalues.
3. Schur’s Lemma: $A \in \mathbb{R}^{n \times n}, U^{-1}AU = T$, $T$ is upper triangular with the eigenvalues of $A$ along the diagonal and $U$ is orthogonal.
4. When $A$ does not have a complete set of eigenvectors, use the Jordan form.
Section VI: Pos. Def. Matrices & Norms

1. Positive Definite Matrices
2. Tests for Positive Definiteness
3. Vector and Matrix norms
4. Norms and Condition Number
5. The Singular Value Decomposition (SVD)
6. Applications of the SVD

Positive Definite Matrices

1. Consider the problem of identifying the minimum of two functions: \( F(x, y) = 7 + 2(x + y)^2 - y \sin y - x^3 \).
   \[ f(x, y) = 2x^2 + 4xy + y^2 \]
2. We shall investigate whether either \( F \) or \( f \) have a minimum at \( x = y = 0 \).
Positive Definite Matrices: II

1. The first-order conditions to check on stationarity are as follows:

\[ \frac{\partial F}{\partial x} = 4(x + y) - 3x^2 = 0 \]
\[ \frac{\partial F}{\partial y} = 4(x + y) - y \cos y - \sin y = 0 \]
\[ \frac{\partial f}{\partial x} = 4(x + y) = 0 \]
\[ \frac{\partial f}{\partial y} = 4x + 2y = 0 \]

2. Therefore \((x, y) = (0, 0)\) is a stationary point for \(f\) and \(F\), the question now is whether it is a minimum?

Positive Definite Matrices: III

1. The second-order conditions:

\[ \frac{\partial^2 F}{\partial x^2} = 4 - 6x = 4 \]
\[ \frac{\partial^2 F}{\partial x \partial y} = 4 \]
\[ \frac{\partial^2 F}{\partial y^2} = 4 + y \sin y - 2 \cos y = 4 \]
\[ \frac{\partial^2 f}{\partial x^2} = 4 \]
\[ \frac{\partial^2 f}{\partial x \partial y} = 4 \]
\[ \frac{\partial^2 f}{\partial y^2} = 2 \]
Positive Definite Matrices: IV

1. Therefore the $F$ and $f$ behave in a very similar fashion near the origin.

2. The higher degree terms have no impact on the question of a local minimum but they make a difference to a global minimum; for instance a term like $-x^3$ may pull $F$ to $-\infty$.

3. A quadratic form, on the other hand, has no higher order terms: $a x^2 + 2bxy + cy^2$ has a stationary point at the origin. Moreover if it is a local minimum, then it is a global minimum.

4. If $f$ is strictly positive at all other points and zero at $(0,0)$ then $f$ is called positive definite.

Positive Definite Matrices: V

1. For a function $F$ in two variables, we have three second derivatives $F_{xx}, F_{xy} = F_{yx}$ and $F_{yy}$. These three derivatives specify a quadratic function $f$ and will also define whether $f$ has a minimum or not. So the question remains as to what are the conditions on $a, b$ and $c$ for a general quadratic so as to ensure that $f$ is positive definite. Note that this only ensures that $F$ has a local minimum.

2. Let us now identify some necessary conditions for positive (negative) definiteness: If $f$ is positive definite then:
   
   (a) $a > (\triangle) 0 (\frac{\partial^2 f}{\partial x^2} > 0)$.
   
   (b) $ac > b^2 \geq 0$.
Positive Definite Matrices: VI

1. If $ac < b^2$, then $f$ is indefinite and the stationary point is a saddle point: e.g. $f = 2xy$.
2. We shall now see where linear algebra fits in.

A Matrix Approach

1. Consider the quadratic form discussed in the previous section:
   
   $$ax^2 + 2bxy + c^2 = [x \ y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

2. When we extrapolate this to the case of $n$ variables:
   
   $$x^TAx = [x_1 \ x_2 \dots \ x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

3. How does one now decide whether the stationary point is a maximum or a minimum? $A$ is the second derivative matrix. Note that both pure quadratics and general functions will have a second-derivative or a Hessian matrix.
A Matrix Approach: II

1. \( F(x) = F(0) + x^T(\nabla F) + \frac{1}{2}x^TAx + \) higher order terms. The second term is zero because of the first order conditions and the second term will decide whether \( F \) is a minimum not.

2. This leads us to the next section which discusses tests for positive definiteness.

Tests for Positive Definiteness

1. **A is positive definite if** \( x^TAx > 0 \) **for all nonzero vectors** \( x \).

2. A real symmetric matrix is positive definite if and only if
   (a) \( x^TAx > 0 \) for all nonzero vectors \( x \).
   (b) \( \lambda(A) > 0 \).
   (c) All upper left submatrices \( A_k \) have positive determinants.
   (d) All pivots (without row exchanges) satisfy \( d_i > 0 \).
Tests for Positive Definiteness: II

1. A is symmetric positive definite if and only if there is a matrix $R$ with independent columns, such that $A = R^T R$. (Remember that $x^T A x = x^T R^T R x = ||Rx||^2 > 0$). One choice for $R$ is the upper triangular matrix and this is called the **Cholesky Factorization**.

2. Finally we shall try and get a physical appreciation of the quadratic form through an example: $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ and $x^T A x = 5u^2 + 8uv + 5v^2 = 1$. This represents an ellipse in the $u-v$ plane centered at (0,0).

Tests for Positive Definiteness: III

1. The axes point in the direction of the eigenvectors of $A$ and are orthogonal because the matrix is symmetric. The major axis corresponds to the smallest eigenvalue and the minor axis corresponds to the largest eigenvalue of $A$. The eigenvalues are identified as 1 and 9.

2. The ellipse equation can be rewritten as

$$5u^2 + 8uv + 5v^2 = \left(\frac{u}{\sqrt{2}} - \frac{v}{\sqrt{2}}\right)^2 + 9\left(\frac{u}{\sqrt{2}} + \frac{v}{\sqrt{2}}\right)^2 = 1.$$
Vector and Matrix Norms

1. We shall now review some of the matrix norms used in numerical linear algebra:

2. One should recall the basic vector norms:

   \[ ||x||_p = \left( |x_1|^p + \ldots + |x_n|^p \right)^{\frac{1}{p}}, \quad p \geq 1. \]
   \[ ||x||_1 = (|x_1| + \ldots + |x_n|). \]
   \[ ||x||_2 = (|x_1|^2 + \ldots + |x_n|^2)^{\frac{1}{2}}. \]
   \[ ||x||_\infty = \max_{1 \leq i \leq n} |x_i|. \]

3. Frobenius Norm: \[ ||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2}. \]

Vector and Matrix Norms: II

1. Some important matrix norms:

   \[ ||A||_p = \max_{x \neq 0} \frac{||Ax||_p}{||x||_p}. \]
   \[ ||A||_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|. \]
   \[ ||A||_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|. \]
Norm and Condition Number

1. Consider a perturbation of the following system:

\[ A(x + \delta x) = b + \delta b. \]
\[ A(\delta x) = \delta b. \]
\[ \delta x = A^{-1} \delta b. \]

2. There will be a large change in the solution if \( A^{-1} \) is large or \( A \) is nearly singular. Note that singularity of \( A \) implies that at least one of its eigenvalues is zero.

3. The eigenvalues of \( A \) are defined as \( 0 < \lambda_1 \leq \lambda_2 \ldots \leq \lambda_n \) and the corresponding eigenvectors are \( x_1, x_2, \ldots x_k \).

4. Any vector \( \delta b \) has to be a linear combination of the eigenvectors (why?) and the worst error is when \( \delta b \) is in the direction of the first eigenvector \( x_1 \).

Norm and Condition Number: II

1. If \( \delta b = \epsilon x_1 \) then \( \delta x = \frac{\delta b}{\lambda_1} \).

2. Thus the amplification is the greatest when \( \lambda_1 \) is closest to zero or \( A \) is nearly singular.

3. Therefore, the error of size \( ||\delta b|| \) is amplified by \( \frac{1}{\lambda_1} \) which is the largest eigenvalue of \( A^{-1} \). The relative error normalizes the problem: \( \frac{||\delta x||}{||x||} \).

4. The condition number of \( A \) is given by \( c = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \). It should be reiterated that the a matrix could be nearly singular yet well conditioned: e.g. \( A = 1/10 \), which has determinant \( 10^{-n} \) but has condition number = 1.
Norm and Condition Number: III

1. The norm of $A$ is defined by $||A|| = \max_{x \neq 0} \frac{||Ax||}{||x||}$. In other words, $||A||$ is an upper bound of the amplification power of $A$. This is clear from $||Ax|| \leq ||A|| ||x||$.

2. Finally, the condition number of $A$ can also be represented as $c = ||A|| ||A^{-1}||$.

Singular Value Decomposition

1. Earlier in the course, we had discussed the $LU$ decomposition and the $QR$ orthogonalization. In this section, we shall discuss the Singular Value Decomposition (SVD).

2. Recall that a symmetric matrix can be factored as follows: $A = Q \Lambda Q^T$ where $Q$ is an orthonormal eigenvector matrix and $\Lambda$ is a diagonal matrix with the eigenvalues.

3. **Singular Value Decomposition**: Any $m$ by $n$ matrix $A$ can be factored into $A = U \Sigma V^T = (\text{orth})(\text{diag})(\text{orth})$. The columns of $U$ ($m$ by $m$) are the eigenvectors of $AA^T$ and the columns of $V$ ($n$ by $n$) are eigenvectors of $A^T A$. 
Singular Value Decomposition: II

1. The rank of $A$ is given by the number of nonzero singular values of $A$. In fact we can go so far as to say that:
   (a) $\mathcal{R}(A) = \text{span}(u_1, u_2, \ldots, u_r)$.
   (b) $\mathcal{N}(A^T) = \text{span}(u_{r+1}, u_{r+2}, \ldots, u_m)$.
   (c) $\mathcal{R}(A^T) = \text{span}(v_1, v_2, \ldots, v_r)$.
   (d) $\mathcal{N}(A) = \text{span}(v_{r+1}, v_{r+2}, \ldots, v_n)$.

2. In other words, the SVD reveals the four fundamental subspaces.

3. Remember that $A = U\Sigma V^T$ can be rewritten as $AV = U\Sigma$ implying that multiplying $A$ by a column of $V$ results in a column of $U$.

Singular Value Decomposition: III

1. What about the singular values?
   $$AA^T = (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma\Sigma^T U^T$$ and
   $$A^TA = (V\Sigma^T U^T)(U\Sigma V^T) = V\Sigma^T\Sigma V^T$$

2. $\implies$ that $U$ must be the eigenvector matrix for $AA^T$ and the eigenvalue matrix must be $\Sigma\Sigma^T$. Similarly, $V$ must be the eigenvector matrix for $A^TA$ and the eigenvalue matrix must be $\Sigma\Sigma^T$. 
Singular Value Decomposition: IV

1. The $r$ singular values on the diagonal of $\Sigma$ ($m$ by $n$) are the square roots of the nonzero eigenvalues of both $AA^T$ and $A^TA$.

2. We shall now provide an example of the Singular Value Decomposition:

$$A = \begin{bmatrix}
-1 \\
2 \\
2
\end{bmatrix} = \begin{bmatrix}
-1/3 \\
2/3 \\
2/3
\end{bmatrix} \begin{bmatrix}
2/3 \\
-1/3 \\
2/3
\end{bmatrix} \begin{bmatrix}
3 \\
0 \\
0
\end{bmatrix} [1].$$

3. This implies that $A^TA$ is 1 by 1 while $AA^T$ is 3 by 3. They both have eigenvalues $9$ ($3^2$). In fact, $A^TA$ has rank 1.

Applications of the SVD

1. SVD is a numerically stable operation since $\|Qx\|^2 = x^TQ^TQx = \|x\|^2$ and multiplication by $Q$ does not destroy the scaling.

2. There are many applications of the SVD:
   (a) Image Processing.
   (b) Determining the rank of a matrix $A$ by checking the rank of $A^TA$.
   (c) Least Squares (shall be discussed)

3. Recall that we found the least squares solution to rectangular system through a system of normal equations. In particular $\tilde{x} = (A^TA)^{-1}(A^Tb)$, required that $A$ had to have full rank. If not, we could not solve the systems of normal equations.
Applications of the SVD: II

1. If \( A \) had dependent rows, then the equations may have no solution. In such a case, the vector \( b \) could be projected onto the column space of \( A \) and the system \( A\bar{x} = p \) will need to be solved. This is solvable since \( p \) is in the column space of \( A \).

2. If \( A \) has dependent columns, the solution \( \bar{x} \) will not be unique and we could choose a minimum norm solution. We shall call this solution \( x^+ \).

3. We shall present an example in which both forms of rank-deficiency exist and use the example to demonstrate the utility of the SVD.

\[
A = \begin{bmatrix}
\sigma_1 & 0 & 0 & 0 \\
0 & \sigma_2 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\quad b = \begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\]

Applications of the SVD: III

1. This equation is important because it shows how one obtains the solution \( x^+ \) from a matrix called the pseudo-inverse of \( A \) and is denoted by \( A^+ \).

2. Some properties of the pseudo-inverse and the min-norm solution:
   (a) If \( A \) is invertible then \( A^+ = A^{-1} \).
   (b) \( (A^+)^+ = A \).
   (c) \( x^+ \) is always in the row space of \( A \).

3. Now for a general matrix \( A \), the pseudoinverse can be calculated from the SVD. Therefore we have \( A = UV^T \) and \( A^+ = V\Sigma^+U^T \). Moreover the minimum length solution to \( Ax = b \) is given by \( x^+ = A^+b = V\Sigma^+U^Tb \).
Section VI: Review

1. $A$ is symmetric positive definite $\iff A = R^T R$, where $R$ is nonsingular. If $R$ is upper triangular, then $A = R^T R$ is called the Cholesky factorization and $R$ is the Cholesky factor.

2. A real matrix $A$ is SPD (SPSD) $\iff$
   (a) $x^T A x > (\geq) 0, \forall x \neq 0$.
   (b) $\lambda_i(A) > (\geq) 0$.

3. The condition number of $A$ is given by $c = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}$ and gives a bound on the relative error in $x$ when there is a perturbation in $b$.

4. The norm of $A$ is given by $||A|| = \max_{x \neq 0} \frac{||Ax||}{||x||}$ and $||Ax|| \leq ||A|| ||x||$.

5. Condition number of $A$ is also given by $c = ||A|| ||A^{-1}||$.

Section VI: Review: II

1. Third important factorization: SVD.

2. Singular Value Decomposition: Any $m$ by $n$ matrix $A$ can be factored into $A = U \Sigma V^T = (\text{orth})(\text{diag})(\text{orth})$. The columns of $U$ ($m$ by $m$) are the eigenvectors of $AA^T$ and the columns of $V$ ($n$ by $n$) are eigenvectors of $A^T A$.

3. The rank of $A$ is given by the number of nonzero singular values of $A$. In fact we can go so far as to say that :
   (a) $\mathcal{R}(A) = \text{span}(u_1, u_2, \ldots u_r)$.
   (b) $\mathcal{N}(A^T) = \text{span}(u_{r+1}, u_{r+2}, \ldots u_m)$.
   (c) $\mathcal{R}(A^T) = \text{span}(v_1, v_2, \ldots v_r)$.
   (d) $\mathcal{N}(A) = \text{span}(v_{r+1}, v_{r+2}, \ldots v_n)$. 
Section VI: Review: III

1. In other words, the SVD reveals the four fundamental subspaces.

2. What about the singular values?
   \[ AA^T = (U \Sigma V^T)(V \Sigma^T U^T) = U \Sigma \Sigma^T U^T \] and
   \[ A^T A = (V \Sigma^T U^T)(U \Sigma V^T) = V \Sigma^T \Sigma V^T \]

3. \[ \implies \] that \( U \) must be the eigenvector matrix for \( A A^T \)
   and the eigenvalue matrix must be \( \Sigma \Sigma^T \). Similarly, \( V \)
   must be the eigenvector matrix for \( A^T A \) and the eigenvalue matrix must be \( \Sigma \Sigma^T \).

Section VII: Linear Programming

1. Linear Programming: An introduction
2. The Simplex Method
3. Basic Feasible Solutions
4. Changes of Basis and the Simplex Tableau
5. The Revised Simplex
**Linear Programming**

1. To date, we have been studying the foundations of linear algebra. However, to get a better understanding we shall discuss one of the most oft-used algorithms - The Simplex method for linear programming.

2. Mathematical programming concerns the minimization of a function possibly subject to some constraints. Linear programming, as the name suggests, restricts itself to linear cost functions and linear constraints.

3. An example of an actual problem that can be posed as an LP is stated: Suppose GM makes a profit of $100 on each Chevrolet, $200 on each Buick and $400 on each Cadillac. They get 20, 17 and 14 miles per gallon and Congress insists the average car must get 18 miles per gallon. The plant can assemble a Chevrolet in 1 minute, a Buick in 2 minutes and a Cadillac in 3 minutes. What is the maximum profit in an 8 hour day?

**Linear Programming: II**

1. The cost function can be written as a profit maximization function while the constraints are that the average fuel consumption should be 18 miles per gallon and the total number of cars produced is constrained by the 8 hour limitation. Finally, the number of cars of any type should be zero or greater.

2. This can be stated as:

\[
\begin{align*}
\text{max} & \quad 100x + 200y + 400z \\
\text{s.t.} & \quad 20x + 17y + 14z \geq 18(x + y + z) \\
& \quad x + 2y + 3z \leq 480. \\
& \quad x, y, z \geq 0.
\end{align*}
\]
The Simplex Method

1. This section will discuss the formulation of a linear program with \( n \) unknowns and \( m \) constraints. The linear cost function and constraint system lends itself to a linear algebraic formulation.

2. We are given:
   
   (a) An \( m \) by \( n \) matrix \( A \).
   
   (b) A column vector \( b \) with \( m \) components.
   
   (c) A column vector \( c \) with \( n \) components.

3. The vector \( x \) has to be "feasible" \( \implies \) it has to satisfy the constraints \( Ax \geq b \) and \( x \geq 0 \). The "optimal"vector is one with lowest cost, cost being \( c_1 x_1 + \ldots c_n x_n \).

The Simplex Method: II

1. The formulation of the linear program can thus be stated as

   \[
   \min \ c^T x \\
   Ax \geq b, \\
   x \geq 0.
   \]

2. What is the geometrical interpretation of this problem? Let us discuss the constraint system first: Each inequality constraint results in a halfspace. The feasible region is common to the \( n \) halfspaces and is given by the positive quadrant in \( n \) dimensional space.
The Simplex Method:  III

1. $Ax \geq b$: Again each of the $m$ equations refers to a halfspace and the feasible region refers to the intersection of these $m$ halfspaces. Taken along with the non-negativity constraints, the feasible region is the intersection of $m + n$ constraints.

2. The cost function $c^T x = k$ represents a plane in $n$ space for each value of $k$. As $k$ varies between $-\infty$ and $+\infty$, it spans across the entire $n$-space. The optimal $x^*$ is the point or set of points where the cost function plane first touches the feasible region. Remember we are assuming that the feasible region is bounded from below, else the optimal cost is $-\infty$.

3. To recap, we have a $n$-dimensional polyhedron and we would like to determine at which point a plane touches this polyhedron such that the cost associated with such a point is minimized.

Geometry of the Simplex Method

1. The basic idea behind the simplex method is quite simple:
   (a) Find a vertex of the feasible region (polyhedral) (Phase I).
   (b) Keep moving from vertex to vertex of the feasible set to obtain the optimal solution such that the cost reduces in each step (if it exists) or determine that cost is unbounded below (Phase II).
   (c) The optimal solution will be given by a vertex from which all the edges result in increased cost.
Linear Algebraic Representation

1. In this section, we shall discuss the linear algebra associated with the Simplex method.

2. A vertex or a corner represents an intersection of \( n \) planes. Note that the feasible region is determined by \( m \) inequalities of the form \( Ax \geq b \) and \( n \) in equalities of the form \( x \geq 0 \). Therefore, a vertex is the result of \( n \) of these being equalities while satisfying the other \( m \) constraints. This also tells us that the maximum possible number of intersections is \( \frac{(n+m)!}{n!m!} \) which is the number of possibilities when choosing \( n \) planes from \( n + m \) possible planes.

Linear Algebraic Representation: II

1. Consider what happens if one removes a plane from the \( n \) planes that make up a vertex. We are left with \( n - 1 \) equations in \( n \) unknowns - the system is underdetermined and there is a single degree of freedom. There are \( n \) edges associated with any vertex and movement along one of these edges is the decision taken in phase II.

2. We shall rewrite the system \( Ax \geq b \) as \( Ax - Iw = b \) where \( w \) is a vector of slack variables. The reason for this reformulation is to express the system as a set of equality constraints. Remember that \( x \in \mathbb{R}^n, w \in \mathbb{R}^m \). The matrix form of this system is

\[
\begin{bmatrix}
A & -I
\end{bmatrix}
\begin{bmatrix}
x \\
w
\end{bmatrix} = b.
\]
Linear Algebraic Representation: III

1. We shall now denote $A$ as the matrix $[A - I]$ and $x$ as $[x \, w]^T$ for purposes of simplicity. The resulting system now becomes the following with $A \in \mathbb{R}^{m \times m+n}$ and $x \in \mathbb{R}^{n+m}$:

$$\begin{align*}
\text{Min} & \quad c^T x \\
Ax & = \quad b, \\
x & \geq \quad 0.
\end{align*}$$

2. We are now in a position to begin the Simplex method: Phase I. It may be recalled that this involved the determination of an initial vertex. This is done quite simply by setting the first $n$ components of $x$ to 0. This results in the system $Ax = b$ being a system in $m$ equations and $m$ unknowns $\Rightarrow$ a vertex can be found by solving this set of equations.

Linear Algebraic Representation: IV

1. Basic Variables The $m$ components that shall be defined by the $m$ equations are known as the basic variables

2. Nonbasic Variables The remaining $n$ variables that are set to zero are known as the non-basic variables.
Basic Feasible Solutions

1. The corners or vertices of the feasible set are the basic feasible solutions of $Ax = b$. A solution is basic when $n$ of its $m + n$ components are zero and it is feasible when $x \geq 0$. Phase I finds one basic feasible solution (bfs) and Phase II proceeds to find the optimal, a step at a time.

2. We shall present an example of the same:

$$Ax = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 2 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix} = b.$$

Basic Feasible Solutions: II

1. This example represents the intersection of $2x_1 + x_2 = 6$ and $x_1 + 2x_2 = 6$. The slacks are denoted by $x_3$ and $x_4$. The basic variables are $x_2$ and $x_3$ while the other two are non-basic.

2. What happens now? As we move to an adjacent corner, one of the basic variables will become non-basic and one of the non-basic variables will become basic. Moreover, the variables which stay basic may change in values.

3. The main decision is to decide which variable to drop from the basis and which one to bring into the basis.
Change of basis

1. We shall now provide an example of an entering variable and a leaving variable.

2.

\[
\begin{align*}
\min & \quad 7x_3 - x_4 - 3x_5 \\
& \quad x_1 + 0x_2 + x_3 + 6x_4 + 2x_5 = 8 \\
& \quad 0x_1 + x_2 + x_3 + 0x_4 + 3x_5 = 9
\end{align*}
\]

3. We shall start from a vertex where \( x_1 = 8 \) and \( x_2 = 9 \) (both are basic) and the remainder are non-basic variables. \( x_3 = x_4 = x_4 = 0 \).

4. To decide which variable should enter, consider that the coefficient of \( x_5 \) is the lowest while it would be foolish to allow \( x_3 \) to enter since it would raise the cost. Therefore, we choose \( x_5 \) as the variable to enter.

Change of basis: II

1. Now \( x_5 \) is allowed to become basic in each of the equations (we shall see that it can become basic in only one of them). For \( x_5 \) to become basic in the first equation, \( x_5 \) has to increase all the way to 4 while it has to increase to 3 in the second equation. Since it is increasing from zero, clearly it becomes basic in the second equation first and this decides which variable becomes non-basic: it is \( x_2 \). Therefore the new system can be written as follows:
Change of basis: III

1. To recap, we take the ratios of the coefficients of \( b_i \) and the coefficients of \( x_{enter} \) or 8/2 and 9/3 and compare. We choose the smaller one to decide which basic variable to drop. Moreover, \( x_5 \) in the objective function can be expressed in terms of the second equation such that the cost function is only in terms of the non-basic variables. Similarly the first equation is rewritten so that \( x_5 \) is replaced by the second equation for \( x_5 \).

2. Therefore we have

\[
\begin{align*}
\min 7x_3 - x_4 - 3x_5 &= 7x_3 - x_4 - 3(9 - x_2 - x_3) \\
\min 8x_3 - x_4 + 3x_2 &= 27 \\
x_1 - 2/3x_2 + 1/3x_3 + 6x_4 + 0x_5 &= 2 \\
0x_1 + 1/3x_2 + 1/3x_3 + 0x_4 + x_5 &= 3
\end{align*}
\]

Change of basis: IV

1. Remark I: We should only be considering positive ratios when deciding which variable to drop. For instance if the coefficient of \( x_5 \) was -3, then increasing then increasing \( x_5 \) would require even more of \( x_2 \) to keep the equation consistent. Thus \( x_2 \) would never leave. If all coefficients of \( x_5 \) are negative, \( \implies \) the solution is unbounded since the cost coefficient of \( x_5 \) is negative and it can be increased indefinitely without affecting feasibility.

2. Remark II The next step is taken similarly. We choose \( x_4 \) to enter the basis and the ratios to be compared are 1/3 and 3/0 \( \implies \) that \( x_1 \) will leave the basis. The final vertex is \( x_1 = x_2 = x_3 = 0, x_4 = 1/3, x_5 = 3 \).

3. Remark III: The step where we replace \( x_5 \) in the first equation is essentially pivoting about the second equation for \( x_5 \).
The Simplex Tableau

1. The decisions of entering and leaving variables can be organized into a tableau which is essentially a large matrix.

2. The tableau has dimension $m + 1 \times m + n + 1$ and is written as:

$$
\begin{bmatrix}
A & b \\
c & 0
\end{bmatrix}
$$

3. Suppose the $m$ components of $x$ that are basic are placed in the right hand corner. Therefore the first $m$ columns form a square matrix which will be hereafter termed as $B$ the basic matrix and the last $n$ columns form an $m \times n$ matrix $N$. The cost vector can also be factored into $c = [c_B \ c_N]$. Note that since $x_N = 0$, the cost is given by $c_B x_B$.

---

The Simplex Tableau: II

1. Now the tableau can be rewritten as $T' =$

$$
\begin{bmatrix}
B & N & b \\
c_B & c_N & 0
\end{bmatrix}
$$

2. The basic variables can be obtained explicitly if we proceed with elimination in system on top. This is equivalent to multiplying by $B^{-1}$ though we shall never compute this explicitly. Remember this is a Gauss-Jordan elimination that results in the identity matrix instead of the usual triangular matrix. $T'' =$

$$
\begin{bmatrix}
I & B^{-1} N & B^{-1} b \\
c_B & c_N & 0
\end{bmatrix}
$$

3. Finally, we can now eliminate $c_B$ from the lower system resulting in $T''' =$

$$
\begin{bmatrix}
I & B^{-1} N & B^{-1} b \\
0 & c_N - c_B B^{-1} N & -c_B B^{-1} b
\end{bmatrix}
$$
The Simplex Tableau: III

1. Now that the tableau is ready, we shall interpret each of its rows.

\[ x_B + B^{-1}Nx_N = B^{-1}b, \]
\[ cx = (c_N - c_BB^{-1}N)x_N + c_BB^{-1}b, \]

2. The basic variables are given by \( B^{-1}b \) on the far right of the tableau since \( x_N = 0 \).

3. The current cost is merely the cost of the basic variables, \( c_BB^{-1}b \) since \( x_n = 0 \).

4. Finally, reduced costs are given by \( c_N - c_BB^{-1}N \) and if all the entries are positive, the solution is optimal otherwise one of the variables associated with the negative entries enters the basis. Note that the computation of the reduced costs is called \textbf{pricing out} the variables.

The Revised Simplex

1. We shall now discuss a step of revised simplex, a boiled down version of the simplex.
   (a) Compute the row vector \( \lambda = c_BB^{-1} \) and the reduced costs \( r = c_n - \lambda N \).
   (b) If \( r \geq 0 \) then stop, else pick \( r_i \) as the most negative component of the reduced costs and choose the \( i^{th} \) column of \( N \) to enter the basis. This shall be denoted by \( u \).
   (c) Compute \( v = B^{-1}u \).
   (d) Calculate the ratios of \( B^{-1}b \) to \( B^{-1}u \), admitting only positive components of the latter. If there are no positive components, the optimal cost is \( -\infty \); if the smallest positive ratio occurs at component \( k \), then the \( k^{th} \) column of \( B \) will leave.
   (e) \( B \) is updated and we return to the first step.
The Revised Simplex: II

1. $B^{-1}$ is calculated explicitly at the first vertex and we have the identity matrix on the top left of the tableau. Once $v$ is computed, it replaces the column $k$ of the identity matrix. To recover the identity matrix, we multiply by an elimination matrix which is shown below:

$$
\begin{bmatrix}
1 & v_1 \\
\vdots & \vdots \\
v_k & \\
\vdots & \vdots \\
v_n & 1
\end{bmatrix}^{-1} = 
\begin{bmatrix}
1 & -v_1/v_k \\
\vdots & \vdots \\
-1/v_k & \\
\vdots & \vdots \\
-v_n/v_k & 1
\end{bmatrix}
$$

The Revised Simplex: III

1. Instead of recomputing $B^{-1}$ at each step, the elimination matrices are stored and applied as required. After a fixed number of simplex steps, the explicit inverse is computed and the stored elimination matrices are flushed.

2. In general, it is seen that method only touches $m$ or $1.5m$ different corners implying an operation count of $m^2n$. However, examples can be constructed such that the simplex method will try every corner. The Simplex method is said to perform in an average time that is polynomial.

3. Khachian’s method as well as Karmarkar’s method showed that the linear programming problem could be solved in polynomial time if one stayed inside the feasible set.