Matrix Review

A. Definitions

B. Matrix Operations

C. Determinants

D. Chio’s Method

E. Inversion

F. Linear Dependence

G. Convex Sets

H. Extreme Points

G. Hyperplanes and Halfspaces

H. Rays and Directions

G. Convex Cones and Functions

H. Polyhedral Representations

I. Homework #1
A. Definitions

An \textit{n-vector} or \textit{vector} in $E^n$ is an ordered $n$–tuple $\mathbf{x} = (x_1, \ldots, x_n)$ of real numbers $x_i$ called the components of $\mathbf{x}$. A matrix is a rectangular array of elements; it has no “value”

\[
A = \begin{pmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{pmatrix}
\]

The size of the matrix is given as $(M \times N)$ where $M$ is the number of rows and $N$ is the number of columns.

A $(1 \times N)$ matrix is called a row vector. An $(M \times 1)$ matrix is called a column vector. If $M=N$, the matrix is called a \textit{square} matrix. If all elements are zero it is a \textit{zero} matrix. A square matrix with all diagonal elements equal to one and all other elements zero is called an \textit{identity} matrix.

Transportation Example from previous lecture:

\[
\begin{pmatrix}
  (1, a) & (1, b) & (1, c) & (2, a) & (2, b) & (2, c) & \mathbf{b} \\
  1 & +1 & +1 & +1 & 0 & 0 & 0 & S_1 \\
  2 & 0 & 0 & 0 & +1 & +1 & +1 & S_2 \\
  a & -1 & 0 & 0 & -1 & 0 & 0 & -D_a \\
  b & 0 & -1 & 0 & 0 & -1 & 0 & -D_b \\
  c & 0 & 0 & -1 & 0 & 0 & -1 & -D_c
\end{pmatrix}
\]
Let \( A \) and \( B \) be two matrices, then:

\[
\begin{align*}
A \leq B \ &\text{implies} \ a_{ij} \leq b_{ij} \ \forall i, j \\
A \geq B \ &\text{implies} \ a_{ij} \geq b_{ij} \ \forall i, j \\
A = B \ &\text{implies} \ a_{ij} = b_{ij} \ \forall i, j
\end{align*}
\]

F. Matrix Operations

Transposition: The transpose of \( A \), \( A^t \) of \( A' \) is formed by interchanging the rows and columns of \( A \):

\[
A = \begin{pmatrix}
3 & 1 & 0 \\
5 & -1 & 5 \\
7 & 2 & 7
\end{pmatrix}
\quad A^t = \begin{pmatrix}
3 & 5 & 7 \\
1 & -1 & 2 \\
0 & 5 & 7
\end{pmatrix}
\]

F.1 Addition and Subtraction

Two matrices \( A \) and \( B \) are “conformable” for addition if they are the same size:

\[
C = A \pm B
\]

is formed by \( c_{ij} = a_{ij} \pm b_{ij} \ \forall i, j \) The commutative and associative laws hold for addition and subtraction:

\[
(A \pm B)^t = A^t \pm B^t
\]
F.2 Multiplication

The product \( AB \) is defined if \( A \) and \( B \) are “conformable” for multiplication in that order, i.e. there are as many rows in \( B \) as there are columns in \( A \). In other words, if \( A \) is \( (M\times N) \) and \( B \) is \( (P\times Q) \) then \( N \) must equal \( P \) for \( AB \) to exist. The eventual product of \( AB \) if \( N=P \) will be a matrix of size \( (M\times Q) \), i.e. it will have as many rows as \( A \) and as many columns as \( B \).

F.3 Row–Column Rule

Let \( A \), \( B \) and \( C \) be matrices and consider the product

\[
A_{MN}B_{NP} = C_{MP}
\]

. The elements of \( C \) are given by:

\[
C_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} \quad \forall i, k
\]

For example, let:

\[
A = \begin{pmatrix}
3 & 2 & -1 \\
2 & 4 & 3 \\
-2 & 4 & 1
\end{pmatrix} \quad B = \begin{pmatrix}
3 & 2 \\
2 & -1 \\
4 & -2
\end{pmatrix}
\]

Let’s find \( C = AB \)

\[
C = AB \rightarrow \begin{pmatrix}
9 & 6 \\
26 & -6 \\
6 & -10
\end{pmatrix}
\]
Matrix multiplication via the above algorithm is known to be $O(n^3)$ but this is not optimal, because there is an even faster way to multiply matrices originally due to Strassen which is $O(N^{2.81})$. Strassen’s algorithm relies on a partitioning of the original matrix which then multiplies $2 \times 2$ matrices and combines them which is an application of a divide-and-conquer approach.

*nota bene (n.b.):* The commutative law does not hold for matrix multiplication, i.e. in general $AB \neq BA$ even if both products exist. However, the Associative and Distributive laws do hold.

\[
(AB)C = A(BC)
\]

\[
A(B + C) = AB + AC
\]

**F.4 Special Cases**

If $S$ is a scalar a (1x1) matrix and $A$ is a matrix, then the product $SA$ is defined and the resultant matrix is formed by multiplying each element of $A$ by $S$.

If $A$ is a row vector and $B$ is a column vector, the product $AB$ is called the inner product of the vectors and is a scalar.

\[
AB = \sum_{j=1}^{n} a_{ij} b_{ji} = \text{some scalar value}
\]

**F.5 Division**

This operation is not defined in matrix algebra; however, an equivalent operation will be discussed, namely *Inversion*. 
F.6 Equivalence

Two matrices are equivalent if one can be obtained from the other by simple row and/or column operations. Row (column) operations:

R1(C1): Interchange two rows (columns)
R2(C2): Multiply a row (column) by a scalar
R3(C3): Add one row (column) to another.

None of these operations change the matrix; matrices do not have to be square to be equivalent.

G. Determinants

Determinants are very valuable in computing such objects as Inverses, Eigenvalues, Positive Definiteness, Solving Equations and many other facets of our study.

If $A$ is a square matrix, its determinant is denoted either by

$$ | A | \text{ or else } det \ A $$

Some of the properties of determinants are:

1) If all the elements of any row or any column are zero, then the value of the determinant is zero.

2) The interchange of any two rows or any two columns of a determinant changes the sign of the determinant, but does not alter its absolute value.
3) If any two rows or any two columns are identical, the determinant $= 0$.

4) If all of the elements of any one row or any one column are multiplied by the same factor, the value of the determinant is multiplied by that factor.

5) The value of a determinant is unchanged by the following operation: Multiply each element of any row by the same factor and add the resulting products, term by term, to the elements of any other row.

For example, multiply the elements of row $i$ by a constant $k$ and add the products to row $h$:

In the resulting determinant, row $i$ is unchanged, but row $h$ has for element:

$$a'_{hj} = a_{hj} + ka_{ij} \quad \text{for each column } j$$

The above row operations can also be performed using the columns; thus column operations are performed as follows:

Multiply each element of any column by the same factor and add the resulting products, term by term, to the elements of any other column.

6) If the product $\mathbf{AB}$ is defined, then:

$$| \mathbf{A} || \mathbf{B} | = | \mathbf{AB} |$$

The determinant of the product is the product of the determinants.
H. Chio’s Method for Evaluating Determinants

**Step 1.0:** See if any element of the determinant $= 1$. If not, use row or column operations to produce such an element.

**Step 2.0:** The element $= 1$ is called the “pivotal” element. The row and column intersecting the pivotal element (row $r$, column $s$) are deleted. Every remaining element $a_{ij}$ in the determinant is diminished by the product of the elements perpendicular to $a_{ij}$ in the eliminated row and column. Thus:

$$a'_{ij} = a_{ij} - a_{is}a_{rj}$$

**Step 3.0:** The entire determinant is multiplied by

$$(-1)^{r+s}$$

**Step 4.0:** When all the elements have been computed, the result is a determinant one order smaller than the previous determinant. Repeat the process until a determinant of order two is obtained and evaluate it in the standard manner:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} (a_{11}a_{22} - a_{21}a_{12})$$
**H.1 Example**

\[
D = \begin{pmatrix} 3 & 6 & 11 \\ 2 & -5 & 6 \\ 8 & 2 & -6 \end{pmatrix} \quad D = \begin{pmatrix} 3 & 6 & 11 \\ 2 & -5 & 6 \\ 4 & 1 & -3 \end{pmatrix}
\]

where, the pivot row \( r = 3 \) and the pivot column \( s = 2 \).

\[
D = 2 \left( -1 \right)^{3+2} \begin{pmatrix} 3 & 6 & 11 \\ 2 & -5 & 6 \\ 4 & 1 & -3 \end{pmatrix} = 2 \left( -1 \right) \left[ (-21 * -9) - (22 * 29) \right] = -2(189 - 638) = -2(-449) = 898
\]

**H.2 Singularity**

The square matrix \( A \) is said to be “singular” if

\[
| A | = 0
\]

A square matrix whose determinant does not vanish is called a \textit{full-rank} or \textit{nonsingular} matrix.
I. Inversion

Inversion is a crucial process in solving sets of equations, Newton’s algorithm and many other routines. It is computationally very expensive.

Since division cannot be performed with matrices, we instead multiply by the “inverse”. The algebraic equivalent is to multiply a number by the reciprocal of another number rather than divide it.

If \( B \) and \( C \) are both square, non–singular matrices, then:

\[
BC = CB = I \text{(the Identity matrix)}
\]

and \( B \) is the inverse of \( C \) denoted by \( C^{-1} \), likewise, \( C \) is the inverse of \( B \). *Only square non–singular* matrices have an inverse, and it is unique.

I.1 Inverse Relationships

For a 2x2 matrix, its inverse can be computed as:

\[
A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ then } A^{-1} = \frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}
\]

The following relationships hold for the inverse as well as the transpose:

\[
(A^{-1})^{-1} = A \quad (AB)^{-1} = B^{-1}A^{-1}
\]
I.2 Computing the Inverse

Determinants can be used to compute the inverse, but this is not a very useful way to find an inverse. The “sweep out” or row transformation method is preferred. To utilize this method, you merely augment the matrix with an identity matrix on the right and partition the matrix, i.e. from $\mathbf{A}$, form the matrix:

$$
(\mathbf{A} \mid \mathbf{I})
$$

Now apply row operations to transform $\mathbf{A}$ to $\mathbf{I}$, in the same process $\mathbf{I}$ will become $\mathbf{A}^{-1}$.

Also, if we augment the column vector $\mathbf{b}$ to the right of $\mathbf{A}$ and proceed above, we will obtain the solution to the equations: $\mathbf{A}\mathbf{x} = \mathbf{b}$.

I.3 Example

Let

$$
\mathbf{A} = \begin{pmatrix}
1 & 2 & 3 \\
-2 & 1 & 5 \\
0 & 1 & 2 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 3 & 1 & 0 & 0 \\
-2 & 1 & 5 & 0 & 1 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 \\
\end{pmatrix}
$$

Multiply row 1 by 2 and add to row 2

$$
\begin{pmatrix}
1 & 2 & 3 & 1 & 0 & 0 \\
0 & 5 & 11 & 2 & 1 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 3 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 \\
0 & 5 & 11 & 2 & 1 & 0 \\
\end{pmatrix}
$$

Interchange row 3 with row 2
\[
\begin{pmatrix}
1 & 2 & 3 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 \\
0 & 5 & 11 & 2 & 1 & 0
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 3 & 1 & 0 & 0 \\
0 & 1 & 2 & 0 & 0 & 1 \\
0 & 0 & 1 & 2 & 1 & -5
\end{pmatrix}
\]

Multiply row 2 by -5 and add to row 3

\[
\begin{pmatrix}
1 & 0 & 0 & 3 & 1 & -7 \\
0 & 1 & 0 & -4 & -2 & 11 \\
0 & 0 & 1 & 2 & 1 & -5
\end{pmatrix}
\]

thus, \( A^{-1} = \begin{pmatrix} 3 & 1 & -7 \\ -4 & -2 & 11 \\ 2 & 1 & -5 \end{pmatrix} \)

Another advantage of this method is that with a little extra computation, the value of the determinant is given when an upper triangular matrix is found (see dotted lines). The determinant is the product of the elements along the diagonal of the upper triangular matrix.

**J. Linear Dependence**

Dependence and Independence of vectors is crucial in the Kuhn–Tucker conditions as well as some of the algorithms.

A set of vectors \( a_1, a_2, \ldots, a_n \) are linearly independent if they span the space \( E^n \), i.e. has only the solution when \( x_j = 0 \ \forall \ j \).

\[
\sum_{j=1}^{n} x_j a_j = 0
\]

The easiest way to check this result is to see if the equations formed above have a unique solution ( see if \( | A | = 0 \) ). If the determinant vanishes, then they are linearly dependent.
J.1 Example

\[
\begin{align*}
    \mathbf{a}_1 &= \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} \\
    \mathbf{a}_2 &= \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} \\
    \mathbf{a}_3 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}
\end{align*}
\]

\[
\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)
\]

\[
| \mathbf{A} | = \text{det} \begin{pmatrix} 4 & 3 & 1 \\ 3 & 3 & 1 \\ 3 & 4 & 1 \end{pmatrix} = -1
\]

Thus, these three vectors form a three dimensional Euclidean space.

If the vectors span the space, then, any point in the space can be represented as a linear combination of the vectors. Example using the above vectors is to express \( \mathbf{b} = (6, 2, -5)^t \) as a linear combination of the \( \mathbf{a}_j \) vectors:

\[
\begin{align*}
    x_1 \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 6 \\ 2 \\ -5 \end{pmatrix}
\end{align*}
\]

This is simply the solution to the equations which can be obtained by the prior method (or any other) as:

\[
    x_1 = 4 \quad x_2 = -7 \quad x_3 = 11
\]
G. Convex Sets

A set $X$ in $E^n$ is called a convex set if given any two points $x_1$ and $x_2$ in $X$, then $\lambda x_1 + (1 - \lambda) x_2 \in X$ for each $\lambda \in [0, 1]$

H. Extreme Points

$x \in X$ is an extreme point of $X$ if it cannot be represented as a strict convex combination of two distinct points in $X$
G. Hyperplanes and Halfspaces

$H$ in $E^n$ is a set of the form

$$\{x : px = k\}$$

where $p$ is a nonzero vector in $E^n$ and $k$ is a scalar. $p$ is the normal or gradient of $H$.

**Halfspaces:** $H$ divides $E^n$ into 2 regions, called halfspaces:
H. Rays and Directions

A ray is a collection of points of the form

\[ \{x_0 + \lambda d : \lambda \geq 0\} \]

where \( d \) is a nonzero vector. \( x_0 \) is the vertex of the ray and \( d \) is its direction.

Directions of a Convex Set

Given a convex set, a nonzero vector \( d \) is called a (recession) direction of the set, if for each \( x_0 \) in the set the ray \( \{x_0 + \lambda d : \lambda \geq 0\} \) also belongs to the set.

If the set \( X \) is bounded, it has no directions
G. Convex Cones and Functions

A convex cone $C$ is a convex set with the additional property that $\lambda x \in C$ for each $x \in C$ and for each $\lambda \geq 0$. Also, if we have a set of vectors $\{a_j : j = 1, \ldots, k\}$ which generate the cone:

$$C = \left\{ \sum_{j=1}^{k} \lambda_j a_j : \lambda_j \geq 0, \forall j \right\}$$

Convex Function A function $f$ of the vector $(x_1, \ldots, x_n)$ is said to be convex if the following inequality holds for any two vectors $x_1$ and $x_2$:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \text{ for all } \lambda \in [0, 1]$$

<table>
<thead>
<tr>
<th>Convex</th>
<th>Concave</th>
<th>Neither</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

17
H. Polyhedral Representations

A polyhedral set is the intersection of a finite number of halfspaces

$$\{x : Ax \leq b\}$$

$$a^i x \leq b_i, \forall i = 1, 2, \ldots, m$$

Faces, Edges, Adjacent Extreme Points
**Representation Theorem** Let \( X = \{ x : Ax \leq b \} \) represent the feasible region of an LP. Let \( V \) be the set of extreme points \( (x_1, \ldots, x_n) \). If \( X \) is nonempty, then \( V \) is nonempty, and every feasible point \( x \in X \) can be written as:

\[
x = \sum_{j=1}^{k} \lambda_j x_j + \sum_{j=1}^{\ell} \mu_j d_j
\]

\[
\sum_{j=1}^{k} \lambda_j = 1, \lambda_j \geq 0 \ \forall j, \bar{\mu}_j \geq 0, \ \forall j
\]

**Bounded Problem**
Unbounded Problem

Homework #1

Problems from Chapter 1 and 2 of the textbook, please see attached sheets.