MA 421 Prerequisite Review - Linear Algebra

This document is written to be a review of the linear algebra relevant to linear programming and optimization rather than a full linear algebra lecture note. Linear Algebra is the *most* important tool needed to properly understand and appreciate Optimization, so a good background in Linear Algebra is essential. Please reach out if you spot any typos in this document.

1 Systems of Linear Equations

We start our linear algebra review with a brief review of linear equations. Systems of equations appear very frequently throughout linear programming and optimization, so it is important to be very comfortable with matrix operations.

A system of m linear equations in n variables can be written in the form

$$Ax = b,$$

where

$$A \in \mathbb{R}^{m\times n}, \qquad x \in \mathbb{R}^n, \qquad b \in \mathbb{R}^m.$$

Here, A is the matrix containing the coefficients of the system, x is the vector of unknown variables, and b is the right-hand side vector.

For example, consider the system

$$\begin{aligned} x_1 + 2x_2 &= 5, \\ 3x_1 + 4x_2 &= 11. \end{aligned}$$

This can be written in matrix form as

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix}.$$

Solving Systems

One standard method for solving systems of linear equations is *Gaussian elimination.*

To perform Gaussian elimination, we work with the augmented matrix

$$\left[\begin{array}{c|c} A & b \end{array}\right].$$

Recall that the three elementary row operations are:

1. Interchange two rows.
2. Multiply a row by a nonzero scalar.
3. Add a multiple of one row to another row.

These operations do not change the solution set of the system.

For the example above, the augmented matrix is

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 11 \end{array}\right].$$

Performing row reduction gives us

$$\left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 11 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 0 & -2 & -4 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right].$$

Therefore, we get

$$x_1 = 1, \qquad x_2 = 2.$$

A system of linear equations can have one of three possible types of solution sets:

- a unique solution;
- infinitely many solutions;
- no solution

A system is called *consistent* if it has at least one solution, and *inconsistent* if it has no solution

For example, if row reduction produces a row of the form

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \mid c \end{bmatrix}, \qquad c \neq 0,$$

then the system is inconsistent and no solution exists.

If the system is consistent and contains one or more free variables, then the system has infinitely many solutions.

These computations will occur frequently when we study linear programming methods, so it is important to be comfortable with matrix computations

2 Linear Combinations, Span, and Linear Independence

We next review linear combinations, span, and linear independence.

Linear Combinations and Span

Let us consider the vectors

$$v_1, v_2, \ldots, v_k \in \mathbb{R}^n.$$

A *linear combination* of these vectors is any vector of the form

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k,$$

where

$$c_1, c_2, \ldots, c_k \in \mathbb{R}.$$

The set of all possible linear combinations of $v_1, \dots, v_k$ is called their *span*, written as

$$\operatorname{span}\{v_1, \dots, v_k\}.$$

For example, let

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Any vector in $\mathbb{R}^2$ can be written as

$$x_1 v_1 + x_2 v_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Therefore,

$$\operatorname{span}\{v_1, v_2\} = \mathbb{R}^2.$$

Linear Independence

The vectors $v_1, \dots, v_k$ are linearly independent if the equation

$$c_1 v_1 + \cdots + c_k v_k = 0$$

has only the trivial solution

$$c_1 = \cdots = c_k = 0.$$

If there exists a nonzero choice of coefficients satisfying

$$c_1 v_1 + \cdots + c_k v_k = 0,$$

then the vectors are linearly dependent.

For example, consider

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad v_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Since

$$v_2 = 2v_1,$$

the vectors are linearly dependent.

On the other hand,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are linearly independent.

Linear Independence and Matrices

Suppose

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix},$$

where $a_1, \ldots, a_n$ are the columns of A.

The matrix and vector product

$$Ax$$

can be written as

$$Ax = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n.$$

Thus, multiplying a matrix by a vector produces a linear combination of the columns of the matrix.

In particular,

$$Ax = 0$$

asks whether there is a linear combination of the columns of A that equals the zero vector. Therefore, the columns of A are linearly independent exactly when

$$Ax = 0$$

has only the solution

$$x = 0.$$

3 Range, Null Space, and Rank

We now review three closely related ideas associated with a matrix: the range, the null space, and the rank.

Let

$$A \in \mathbb{R}^{m \times n}.$$

Range

The *range* of A is the set of all vectors that can be written in the form

$$Ax$$

for some $x \in \mathbb{R}^n$.

In other words,

$$\boxed{\operatorname{range}(A) = \{Ax : x \in \mathbb{R}^n\}.}$$

Since we know that

$$Ax = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n,$$

where $a_1, \ldots, a_n$ are the columns of A, the range of A is also the span of the columns of A:

$$\operatorname{range}(A) = \operatorname{span}\{a_1, \ldots, a_n\}.$$

For example, consider

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

The second column is twice the first column, so

$$\operatorname{range}(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}.$$

The range is useful when studying the system

$$Ax = b.$$

In particular, the system has a solution exactly when

$$b \in \operatorname{range}(A).$$

Null Space

The *null space*, also called the *kernel*, of A is the set of all vectors x satisfying

$$Ax = 0.$$

$$\text{null}(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

For the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix},$$

we solve

$$Ax = 0.$$

This gives

$$x_1 + 2x_2 = 0,$$

so

$$x_1 = -2x_2.$$

Therefore,

$$\text{null}(A) = \text{span}\left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}.$$

Notice that the null space is closely related to linear independence. The columns of A are linearly independent exactly when

$$\text{null}(A) = \{0\}.$$

Rank

The *rank* of a matrix is the dimension of its range

$$\text{rank}(A) = \dim\bigl(\text{range}(A)\bigr).$$

Equivalently, the rank is the maximum number of linearly independent columns of A.

The rank can also be determined using row reduction. The number of pivot columns in a row-echelon form of A is equal to

$$\text{rank}(A).$$

One of the main relationships between the rank and null space of a matrix is the *Rank–Nullity Theorem.*

If

$$A \in \mathbb{R}^{m \times n},$$

then

$$\text{rank}(A) + \dim\bigl(\text{null}(A)\bigr) = n.$$

The quantity

$$\dim\bigl(\text{null}(A)\bigr)$$

is sometimes called the *nullity* of A.

4 Matrix Multiplication

Multiplying matrices will be needed in this course, especially while studying methods for Linear Programming.

Matrix-Vector Multiplication

Suppose

$$A \in \mathbb{R}^{m\times n}$$

and

$$x \in \mathbb{R}^{n}.$$

Then

$$Ax \in \mathbb{R}^{m}.$$

If

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

then

$$Ax = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}.$$

As we saw earlier, Ax can also be viewed as a linear combination of the columns of A.

Matrix-Matrix Multiplication

Suppose

$$A \in \mathbb{R}^{m\times n}, \qquad B \in \mathbb{R}^{n\times p}.$$

Then the product

$$AB$$

is an $m \times p$ matrix.

The (i, j) entry of AB is

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

It is important to remember that matrix multiplication is associative:

$$(AB)C = A(BC),$$

whenever the products are defined.

However, matrix multiplication is generally not commutative:

$$AB \neq BA.$$

In fact, depending on the dimensions of the matrices, one of these products may be defined while the other is not.

The Transpose

If

$$A \in \mathbb{R}^{m \times n},$$

then the *transpose* of A, denoted by A^T, is the $n \times m$ matrix obtained by interchanging the rows and columns of A.

For example,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

has transpose

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Some useful identities to remember are

$$(A^T)^T = A$$

and

$$(AB)^T = B^T A^T.$$

5 Basis & Invertibility

We now briefly review the ideas of a basis and an invertible matrix.

Basis and Dimension

A collection of vectors

$$v_1, \ldots, v_k$$

forms a *basis* for a vector space V if the vectors are linearly independent and span V.

A basis for V is a collection of vectors that

- spans V, and
- is linearly independent.

For example, the vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^2$.

The number of vectors in a basis is called the *dimension* of the vector space. Thus,

$$\dim(\mathbb{R}^n) = n.$$

Similarly, a basis for the range or null space of a matrix gives us a linearly independent set of vectors that spans that subspace.

Invertible Matrices

Let

$$A \in \mathbb{R}^{n \times n}.$$

The matrix A is called *invertible* if there exists a matrix A^{-1} such that

$$AA^{-1} = A^{-1}A = I.$$

If A is invertible, then the system

$$Ax = b$$

has the unique solution

$$x = A^{-1}b.$$

Several concepts we have already reviewed are equivalent for a square matrix.

For $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

- A is invertible.
- $\operatorname{rank}(A) = n$.
- $\operatorname{null}(A) = \{0\}$.
- The columns of A are linearly independent.
- The columns of A form a basis for $\mathbb{R}^n$.
- For every $b \in \mathbb{R}^n$, the system

  $$Ax = b$$

  has a unique solution.

These relationships connect many of the ideas from the previous sections.

The concept of invertibility will occur frequently in linear programming, especially when we study the simplex method in detail.

6 Eigenvalues & Eigenvectors

We finish this review with eigenvalues and eigenvectors. These concepts are useful for understanding how matrices act in special directions and will appear when we discuss nonlinear optimization.

Let

$$A \in \mathbb{R}^{n \times n}$$

A non-zero vector $v \in \mathbb{R}^n$ is called an *eigenvector* of A if there exists a scalar λ such that

$$Av = \lambda v$$

The scalar λ is called the corresponding *eigenvalue*

To find the eigenvalues of A, we solve

$$\det(A - \lambda I) = 0.$$

Once an eigenvalue λ is known, the corresponding eigenvectors are the nonzero solutions of

$$(A - \lambda I)v = 0.$$

Computing Eigenvalues and Eigenvectors

We now do an example to show how to find eigenvalues and eigenvectors of a matrix

Let us consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

To find the eigenvalues, we solve

$$\det(A - \lambda I) = 0.$$

We have

$$A - \lambda I = \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}.$$

Therefore,

$$\det(A - \lambda I) = (2-\lambda)^2 - 1.$$

Setting this equal to zero gives

$$(2-\lambda)^2 - 1 = 0.$$

Factoring,

$$(1-\lambda)(3-\lambda) = 0,$$

so the eigenvalues are

$$\lambda_1 = 1, \qquad \lambda_2 = 3.$$

We now find the corresponding eigenvectors.

For $\lambda_1 = 1$, solve

$$(A - I)v = 0.$$

This gives

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus,

$$v_1 = -v_2,$$

so one corresponding eigenvector is

$$v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

For $\lambda_2 = 3$, solve

$$(A - 3I)v = 0.$$

This gives

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus,

$$v_1 = v_2,$$

so one corresponding eigenvector is

$$v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Therefore, the matrix has eigenvalue–eigenvector pairs

$$\lambda_1 = 1, \qquad v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

and

$$\lambda_2 = 3, \qquad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

For this course, you should be comfortable finding eigenvalues using

$$\det(A - \lambda I) = 0$$

and then finding the corresponding eigenvectors by solving

$$(A - \lambda I)v = 0.$$