Math 4A Notes

Written by Victoria Kala
vtkala@math.ucsb.edu
Last updated June 11, 2017

Systems of Linear Equations

A **linear equation** is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

where $x_1, ..., x_n$ are variables, $a_1, ..., a_n$ are constants (these are also called the **coefficients**), and b is also a constant.

Example. Consider the following equations:

(a) $4x + 3z = 1$ (b) $8\sqrt{x_1} - 2x_2 + 4x_3^{-1} = 2$ (c) $x^2 + xy + y^2 = 1$ (d) $x_1 + 2x_2 + 3x_3 - 4x_4 = 0$

The equations in (a) and (d) are linear because they can be written in the form above. The equations (b) and (c) are not linear due to the square root, exponents, and multiplication between variables. □

A **system of linear equations** is a collection of one or more linear equations. A system of linear equations is said to be **consistent** if there is a solution, or is said to be **inconsistent** if there is no solution. A consistent system is said to be **independent** if there is exactly one solution, or is said to be **dependent** if there are infinitely many solutions. To summarize:

- Inconsistent: no solution to the system
- Consistent: there is a solution to the system
  - Independent: exactly one solution to the system
  - Dependent: infinitely many solutions to the system

We can write systems of linear equation as an **augmented matrix**. We then use elementary row operations to solve the matrix/system:

- Interchange any two rows
- Multiply by a row by a nonzero constant (scaling)
- Add multiples of rows to each other and replace one of these rows (replacement)

Row Reduction and Echelon Forms

A matrix is in **echelon form** (or **row echelon form**) if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.

2. Each leading entry of a row (called the **pivot**) is in a column to the right of the leading entry of the row above it.

3. All entries in a column below a leading entry (pivot) are zeros.

I personally like to have my pivots (leading entries) be 1, but your textbook does not indicate this preference. The process of getting to echelon form is called Gaussian elimination.

If a matrix is an echelon form, it is said to be in **reduced echelon form** (or **reduced row echelon form**):

1. The leading entry (pivot) in each nonzero row is 1.

2. Each leading 1 is the only nonzero entry in its column.

The process of getting to reduced row echelon form is called Gauss-Jordan elimination.

Example. Consider the following matrices

$$\text{(a)} \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 0 & 3 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 5 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(d)} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \quad \text{(e)} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The matrices (a) and (d) are in row echelon form. The matrices (c) and (e) are in reduced row echelon form. The matrix (b) is neither in echelon nor reduced row echelon form due to the second row. $\square$

If we are solving an augmented matrix, say for example a 3×3 system, then for row echelon our goal will be

$$\left(\begin{array}{ccc|c} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{array}\right) \quad \text{or} \quad \left(\begin{array}{ccc|c} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \end{array}\right)$$

and our goal for reduced row echelon will be

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right).$$

The variables that correspond with pivot columns are said to be **basic variables**. The variables that correspond with columns missing pivots are said to be **free variables**. Whenever there is a presence of a free variable the corresponding system will have infinitely many solutions. In some cases, the presence of a free variable will correspond with a row of zeros (this is especially true for square systems, like a 3×3 system). A system will be inconsistent if we get a false equation such as $0 = 2$.

Vector Equations

A matrix with one column is called a **vector**. A vector with m rows is a vector in the space $\mathbb{R}^m$. Our current operations for vectors right now only consist of addition (including subtraction) and scalar multiplication. We will learn more operations later on.

Let $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$ be vectors. We say that

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n}$$

is a **linear combination** of the vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$, where $c_1, c_2, ..., c_n$ are constants.

We define $\text{span}\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ to be the set of all linear combinations of the vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$. There are infinitely many linear combinations and so $\text{span}\{\}$ is an infinite set.

A vector $\mathbf{b} \in \text{span}\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ if and only if we can write $\mathbf{b}$ as a linear combination of $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$. To do this, we can solve the vector equation

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n} = \mathbf{b}$$

for $c_1, c_2, ..., c_n$ by writing the equation as the augmented matrix

$$\begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} & ... & \mathbf{v_n} & | & \mathbf{b} \end{pmatrix}$$

and reducing the matrix.

The Matrix Equation $A\mathbf{x} = \mathbf{b}$

If A is an $m \times n$ matrix, with columns $\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_n}$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the product of A and $\mathbf{x}$ is given by

$$A\mathbf{x} = \begin{pmatrix} \mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n}.$$

For $\mathbf{b}$ in $\mathbb{R}^m$, the matrix equation $A\mathbf{x} = \mathbf{b}$ has the same solution as the equation

$$x_1\mathbf{a_1} + x_2\mathbf{a_2} + ... + x_n\mathbf{a_n} = \mathbf{b},$$

which also has the same solution as the augmented matrix

$$\begin{pmatrix} \mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} & | & \mathbf{b} \end{pmatrix}.$$

The following is a very useful theorem:

Theorem 1. *Let A be an $m \times n$ matrix. Then the following statements are equivalent (if one is true, all are true; if one is false, all are false):*

(a) For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

(b) Each $\mathbf{b}$ in $\mathbf{R}^m$ is a linear combination of the columns of A.

(c) The columns of A span $\mathbb{R}^m$.

(d) A has a pivot position in every row.

A is a matrix of vectors (the vectors are the columns). If we can show that A has a pivot position in every row using row operations, then the vectors that make up A span the space. We can therefore write any matrix in that space as a linear combination of these vectors.

Solution Sets of Linear Systems

The matrix equation $A\mathbf{x} = \mathbf{0}$ is said to be a **homogeneous** equation. Homogeneous is just a fancy way of saying our equation is equal to "0" (this term will show up in later math courses). $\mathbf{x} = \mathbf{0}$ is certainly a solution to this equation; this is called the **trivial solution**. Some systems have more solutions to this equation, and these are called **nontrivial solutions**. These occur if and only if the equation has at least one free variable.

If a linear system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions, the general solution can be written in the **parametric vector form** (a linear combination of vectors that satisfy the equation).

Example. Suppose a system of equations has the solution $x_1 = 4x_3 - 1, x_2 = -x_3 + 3, x_3 = x_3$ (notice here that x_3 is a free variable). We can write our solution as follows:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4x_3 - 1 \\ -x_3 + 3 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}.$$

This right hand side is in parametric form as it is a linear combination of vectors. □

Linear Independence

A set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ are said to be **linearly independent** if the vector equation

$$c_1\mathbf{v_1} + c_2\mathbf{v_2} + ... + c_n\mathbf{v_n} = \mathbf{0}$$

has the trivial solution $c_1 = c_2 = ... = c_n = 0$. If $c_1, c_2, ..., c_n$ are not all equal to zero and satisfy the above equation, then the vectors are **linearly dependent**. If the vectors are linearly dependent, we say the equation above is the **linear dependence relation** among $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$.

We can also extend linear independence and dependence to a matrix. Consider the matrix $A = \begin{pmatrix} \mathbf{a_1} & \mathbf{a_2} & ... & \mathbf{a_n} \end{pmatrix}$. The columns of a matrix A (remember, columns are vectors) are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution.

Linear Transformations

A **transformation** (also called a function or mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns a vector $\mathbf{x} \in \mathbb{R}^n$ to a vector $T(\mathbf{x}) \in \mathbb{R}^m$. We say a transformation T is **linear** if:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $c \in \mathbb{R}, \mathbf{u}$ in the domain of T.

These properties imply that $T(\mathbf{0}) = \mathbf{0}$ and $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$. Linear combinations are also preserved; that is:

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + ... + c_nT(\mathbf{v}_n).$$

A matrix transformation from $\mathbb{R}^n \to \mathbb{R}^m$ is given by $T(\mathbf{x}) = A\mathbf{x}$ where A is an $m \times n$ matrix. Matrix transformations are linear transformations! The standard matrix representation of A is given by

$$A = \begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \dots & T(\mathbf{e}_n) \end{pmatrix}$$

where $\mathbf{e}_1 = (1, 0, ..., 0), \mathbf{e}_2 = (0, 1, ..., 0), ..., \mathbf{e}_n = (0, 0, ..., 1)$.

Matrix Operations

Recall that the size of a matrix is the number of rows by the number of columns. For example, a 3×100 matrix has 3 rows, 100 columns.

We can perform the following operations with matrices:

- Addition: The sum $A + B$ is defined as long as A and B are the same size, and each entry of $A + B$ is the sum of the corresponding entries in A and B.
- Scalar multiplication: If $c \in \mathbb{R}$ and A is a matrix, then cA is the matrix whose entries are all multiplied by c.
- Matrix Multiplication: The product AB is defined as long the number of columns of A is the same as the number or rows of B. It is helpful to remember the following:

  $$\underbrace{A}_{m\times n} \times \underbrace{B}_{n\times p} = \underbrace{C}_{m\times p}$$

  The ijth entry (that is, the ith row, jth column) of AB is found by multiplying the ith row of A and the jth column of B.
- Transpose: The transpose of a matrix A (denoted by A^T) is found by switching its rows and columns.

Example. Let A be the 3×4 matrix is given by

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 2 \\ 5 & -7 & 0 & 0 \end{pmatrix}.$$

The transpose of A is given by

$$A^T = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 2 \\ 5 & -7 & 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 0 & 5 \\ 2 & 1 & -7 \\ 3 & -1 & 0 \\ 4 & 2 & 0 \end{pmatrix}.$$

Notice that A^T is a 4×3 matrix. (The transpose "swaps" the size of the matrix.)

Some items to remember:

- In general, $AB \neq BA$.
- The cancellation law doesn't always hold; that is, $AB = AC$ does NOT always imply $B = C$.
- $(AB)^T = B^T A^T$

The Inverse of a Matrix

An $n \times n$ matrix A is **invertible** (sometimes called **nonsingular**) if there exists an A^{-1} such that

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

Not all matrices have inverses, however. One way we can check to see if an inverse is invertible is to calculate the determinant (see next section). If $\det A \neq 0$ then A^{-1} exists.

We can use the inverse to solve systems of equations. If A is invertible then the equation $A\mathbf{x} = \mathbf{b}$ has the solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Some nice properties of inverses:

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

We can easily find the inverse of a 2×2 matrix. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

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

provided that $ad - bc \neq 0$.

For larger matrices, we use row operations to find inverses. One way is to use **elementary matrices**. An elementary matrix is the matrix representation of a single row operation and is found by applying a row operation to the identity matrix. For example, if we performed the row operation $-4R1 + R3 \to R3$, then the corresponding elementary matrix would be

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}.$$

If we can find elementary matrices $E_1, ..., E_n$ such that $E_n \cdots E_1 A = I$, then $A^{-1} = E_n \cdots E_1$.

An alternate way is to augment our matrix with the identity matrix:

$$\begin{pmatrix} A & I \end{pmatrix}$$

and then use row operations to reduce to the matrix to

$$\begin{pmatrix} I & A^{-1} \end{pmatrix}.$$

Determinants

One way to find the determinant of a matrix is to use cofactor expansion. We use the following steps:

1. Write down the corresponding sign matrix (checkerboard of $+$ and $-$, starting with $+$).
2. Choose to cross out a row or column (try to pick the one with the most 0's).
3. If you cross out a row, cross out the columns one by one; if you cross out a column, cross out the rows one by one.
4. Calculate the determinant of the remaining matrices. (Note: you may need to repeat the previous steps to do this.)

There is a gnarly formula for this in the textbook, but I think it is better to just look at examples.

Another way is to use row reduction. We can use row operations to write a matrix in row-echelon form (usually upper triangular). The determinant of a triangular matrix is the product of its diagonal entries. The following are the row reduction rules:

1. If you multiply a row by a number k, then you must multiply the determinant by $\frac{1}{k}$.
2. If two rows are interchanged, the determinant changes sign.
3. If a multiple of one row is added to another, there is no change.

A note about Rule (1): I am talking about this as if you are REDUCING the matrix and writing down your row operations as you go. Your book writes this in an alternative way by saying if a row is multiplied by a number then the determinant is multiplied by that number. These are equivalent definitions, so use whichever one you feel more comfortable with.

When to use cofactor expansion vs. row reduction (if not specified in the directions):

- If you are given a matrix with a lot of 0's, use cofactor expansion.
- If you are given a matrix where rows appear to cancel out nicely, use row reduction.
- If you are given a determinant of a matrix and asked to find the determinant of a rearrangement of that matrix, use row reduction.

Otherwise, just use the method you are most comfortable with.

Vector Spaces and Subspaces

A **vector space** is a nonempty set V of objects, called vectors, on which addition and scalar multiplication is defined and has the following properties:

1. If $\mathbf{u}, \mathbf{v} \in V$ then $\mathbf{u} + \mathbf{v} \in V$. (We call this "closed under addition".)
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
4. There is a $\mathbf{0} \in V$ such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$.
5. For each $\mathbf{u} \in V$ there is a $-\mathbf{u} \in V$ such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
6. For each scalar $c \in \mathbb{R}$ and $\mathbf{u} \in V$, $c\mathbf{u} \in V$. (We call this "closed under scalar multiplication.")
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$
9. $c(d\mathbf{u})=(cd)\mathbf{u}$
10. $1 \cdot \mathbf{u}=\mathbf{u}$

$\mathbb{R}^n$ is a vector space. Other vector spaces include polynomials with real coefficients and continuous functions.

A **subspace** of a vector space V is a subset $U \subset V$ with the following properties:

1. $\mathbf{0} \in U$ (note that this is the zero vector of the vector space V)
2. If $\mathbf{u}, \mathbf{v} \in U$ then $\mathbf{u}+\mathbf{v} \in U$. (We call this "closed under addition".)
3. For each scalar $c \in \mathbb{R}$ and $\mathbf{u} \in U$, $c\mathbf{u} \in U$. (We call this "closed under scalar multiplication.")

A subspace is a vector space! Once these three properties are verified, the other properties will follow.
Note that if $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n \in V$, then $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ is a subspace of V.

Basis

Let V be any vector space and $S=\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ be a set of vectors in V. We say that S is a basis for V if the following conditions hold:

(i) S is linearly independent.

(ii) S spans V.

Example. Let $\mathbf{i}=(1,0,0), \mathbf{j}=(0,1,0), \mathbf{k}=(0,0,1)$. The set $S=\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ forms what we call the **standard basis** of $\mathbb{R}^3$. Notice that the set S is linearly independent and this set of vectors spans $\mathbb{R}^3$. (See Midterm 1 Review for information on linear independence and span.)

The Null Space and Column Space

The **null space** of an $m \times n$ matrix A is the set of all solutions to the equation $A\mathbf{x}=\mathbf{0}$. In set notation,

$$Null(A)=\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x}=\mathbf{0}\}.$$

The null space is a subspace of $\mathbb{R}^n$. To show if a vector $\mathbf{v}$ is in the null space of a matrix A, show that $A\mathbf{v} = \mathbf{0}$. If you are asked to find a spanning set for a null space of a matrix A, find the general solution of $A\mathbf{x} = \mathbf{0}$.

The **column space** of an $m \times n$ matrix A is the set of all linear combinations of the columns of A. In set notation,

$$Col(A) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$$

where $A = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n \end{pmatrix}$. The column space forms a subspace of $\mathbb{R}^m$. To show if a vector $\mathbf{v}$ is in the column space of a matrix A, show that $A\mathbf{x} = \mathbf{v}$ has a solution.

The **kernel** and **range** of a linear transformation $T(\mathbf{x}) = A\mathbf{x}$ are just the null space and column of space of A, respectively.

Coordinate Systems and Change of Basis

Since a basis spans a space, then all the vectors in that space can be written as a linear combination of the basis vectors, i.e. for a basis $B = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ of a vector space V, then for all $\mathbf{x} \in V$,

$$\mathbf{x} = c_1\mathbf{v}_1 + ... + c_n\mathbf{v}_n.$$

The vector of the weights $c_1, ..., c_n$ is said to be the coordinate vector of $\mathbf{x}$ with respect to B, i.e.

$$\mathbf{x}_B = \begin{pmatrix} c_1 & c_2 & ... & c_n \end{pmatrix}.$$

We can find a way to change between bases. Let $B = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ and $C = \{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n\}$ be bases of a vector space V. Then there exists an $n \times n$ matrix $P_{B\to C}$ such that

$$\mathbf{x}_c = P_{B\to C}\mathbf{x}_B \quad \text{and} \quad \mathbf{x}_B = P_{C\to B}\mathbf{x}_C$$

where $P^{-1}_{B\to C} = P_{C\to B}$. $P_{B\to C}$ is called the **change of coordinates matrix from B to C**. To find $P_{B\to C}$, set up the augmented matrix

$$\left(\begin{array}{cccc|cccc} \mathbf{c}_1 & \mathbf{c}_n & ... & \mathbf{c}_n & \mathbf{b}_1 & \mathbf{b}_2 & ... & \mathbf{b}_n \end{array} \right)$$

and reduce to get

$$\left(\begin{array}{c|c} I & P_{B\to C} \end{array} \right).$$

Dimension of a Vector Space

The **dimension** of a vector space is the number of elements in its basis. In particular, the dimension of $Null(A)$, or **nullity**, is the number of free variables in the equation $A\mathbf{x} = \mathbf{0}$ and the dimension of $Col(A)$, or **rank**, is the number of pivot columns of A.

If U is a subspace of a vector space V, then $\dim U \leq \dim V$.

Theorem (Rank-Nullity Theorem)**.** *If A is an $m \times n$ matrix then*

$$rank(A) + nullity(A) = n.$$

If A is an $m \times n$ matrix and has rank r, then

dim of	is...
$Col(A)$	r
$Row(A)$	r
$Null(A)$	$n - r$
$Null(A^T)$	$m - r$

Eigenvalues and Eigenvectors

λ is said to be an **eigenvalue** of a square matrix A with nonzero **eigenvector** $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

We can rearrange this equation to get

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{1}$$

Since $\mathbf{x}$ is nonzero, it must be that

$$\det(A - \lambda I) = 0. \tag{2}$$

We use equation (2) to find the eigenvalues of a matrix. After we find the eigenvalues, we then use equation (1) to find their eigenvectors. Equation (2) should yield a polynomial equation, this is sometimes called the **characteristic equation**.

A matrix A is said to be **similar** to a matrix B if there exists an invertible matrix P such that

$$A = PBP^{-1}.$$

Some nice facts:

- The eigenvalues of a triangular or diagonal matrix are the entries on its main diagonal.
- If a matrix has eigenvalue 0 then it is not invertible. (See practice problems.)
- If two matrices are similar then they have the same eigenvalues. (See practice problems.)
- The set of eigenvectors of a matrix (called the **eigenspace**) form a subspace.

Diagonalization

Sometimes we wish to find powers of matrices, like A^{100}. But it would be a very difficult and tedious process to multiply a matrix A out 100 times. To make this process easier we **diagonalize** a matrix: If A is an $n \times n$ matrix and has n distinct eigenvectors, then $A = PDP^{-1}$ where P is the matrix of eigenvectors of A:

$$P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{pmatrix}$$

and D is a diagonal matrix of eigenvalues of A:

$$D = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_m \end{pmatrix}.$$

If $A = PDP^{-1}$ then

$$\begin{aligned} A^k &= A \cdot A \cdots A \\ &= PDP^{-1} \cdot PDP^{-1} \cdots PDP^{-1} \\ &= PD(P^{-1}P)D(P^{-1}P) \cdots (P^{-1}P)DP^{-1} \\ &= PDD \cdots DP^{-1} \\ &= PD^kP^{-1}. \end{aligned}$$

If

$$D = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_m \end{pmatrix}$$

then

$$D^k = \begin{pmatrix} \lambda_1^k & 0 & ... & 0 \\ 0 & \lambda_2^k & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \lambda_m^k \end{pmatrix}.$$

Dot Product, Length, and Orthogonality

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. The **dot product**, or **inner product**, $\mathbf{u} \cdot \mathbf{v}$ is defined to be

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \begin{pmatrix} u_1 & u_2 & ... & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1v_1 + u_2v_2 + ... + u_nv_n.$$

The **norm**, or **length**, of a vector $\mathbf{u} \in \mathbb{R}^n$ is defined by

$$||\mathbf{u}|| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + ... + u_n^2}.$$

Notice that from this definition we have $||\mathbf{u}||^2 = \mathbf{u} \cdot \mathbf{u}$.

The **distance** between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is given by $||\mathbf{u} - \mathbf{v}||$. The vectors $\mathbf{u}$ and $\mathbf{v}$ are said to be **orthogonal** if and only if $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

The dot product of two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is also given by the formula

$$\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||||\mathbf{v}|| \cos\theta$$

where θ is the angle between $\mathbf{u}, \mathbf{v}$. We can rearrange this formula to find the angle between to vectors:

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||||\mathbf{v}||}\right).$$

Properties:

- $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (this says order doesn't matter)
- $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
- $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
- $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$
- $||c\mathbf{v}|| = |c|||\mathbf{v}||$ for any $c \in \mathbb{R}$

Orthogonal Sets

A set of vectors $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ is said to be an **orthogonal set** if each set of vectors is orthogonal, that is, $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ for all $i \neq j$. A basis is said to be an **orthogonal basis** if it is also an orthogonal set.

Let $\{\mathbf{u}_1, ..., \mathbf{u}_n\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y} \in W$,

$$\mathbf{y} = c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n$$

where $c_j = \dfrac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}$, $j = 1, ..., p$. The term

$$\frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}\mathbf{u}_j$$

is said to be the **projection of y onto $\mathbf{u}_j$**.

A **unit vector u**, sometimes called a **normed** vector, is a vector that has length one; that is $||\mathbf{u}|| = 1$. To normalize a vector $\mathbf{v}$, we divide the vector by its length:

$$\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}.$$

An **orthonormal** set is a set of orthogonal unit vectors. A matrix A is said to be **orthogonal** if and only if its column vectors are orthonormal; that is,

$$A^T A = I$$

.

Least Squares Problems

The least squares solution of the system $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}}$ such that

$$||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}||$$

for all $\mathbf{x}$. To find the least squares solution, we solve the system

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

for $\hat{\mathbf{x}}$. The orthogonal projection of $\mathbf{b}$ on the column space of A is the product $A\hat{\mathbf{x}}$.