Math 54 Lecture Notes

Andi Gu

Contents

1 August 22 Lecture

Linear systems are any group of equations in the form $\sum_j a_{ij}x_{ij} = y_i$, where i represents the equation number. However, they can be represented as a matrix mutiplication:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Or, equivalently:

$$A\vec{x} = \vec{y}$$

However, a more concise form is called the 'augmented matrix' form:

$$\left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & y_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & y_m \end{array}\right]$$

We can solve this linear system by performing any one of three operations on the rows of this matrix:

1. Adding a multiple of one row to another
2. Switching rows
3. Multiply a row by a constant

The goal is to reduce the matrix down to row echelon form so that:

1. Any row of 0s is at the bottom
2. The leading entry of a given row is to the right of the one above
3. Every element below a leading entry is 0

For convenience, we define 'pivot positions' as the locations of the leading entries in echelon form.

2 August 24 Lecture

The reduced echelon has leading entries that are all 1s, and everything above and below each leading entry is 0. The reduced echelon form of a linear system is unique, meaning there is one and only form.

There are three types of systems:

1. Inconsistent systems - those with no solutions. Occurs when there is a row of 0s and a non-zero y entry.
2. Those with only one solution. Occurs when there are no 'free' variables.
3. Those with infinite solutions. Occurs when there are free variables.

3 August 27 Lecture

3.1 Brief Introduction on Logic

If we have two phrases A and B, and then we say "If A then B", this means whenever A is true, B is true as well. We can also say A implies B. We can also say A if and only if B, which means if A then B, AS WELL AS if B then A. A and B are either obth true or both false. Saying that "the following are equivalent" means that the following phrases are either all true or all false.

3.2 Vectors

We define $\mathbb{R}$ to be the set of all real numbers. A vector is a matrix with one column:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Now, we can say that $\mathbb{R}^2$ is the set of vectors with two elements, and similarly $\mathbb{R}^n$ is the set of vectors with n entries.

We can now define the linear combination of $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p$ with real number weights $c_1, c_2, \ldots c_p$ as:

$$\vec{y} = \sum_{i=1}^{p} c_i \vec{v}_i$$

An important question arises: we can ask if a given vector $\vec{b}$ can be expressed as a linear combination of some set of vectors.

So the upshot of this is that the equation $\vec{b} = \sum_{i=1}^{p} c_i \vec{v}_i$ has the same solution set as the augmented matrix:

$$\left[\begin{array}{cccc|c} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_p & \vec{b} \end{array}\right]$$

Now, we say that if $\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p$ are vectors in $\mathbb{R}^n$ then their span is:

$$\begin{aligned} span(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p) &= \{\text{linear combinations of} \vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\} \\ &= \{c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots c_p\vec{v}_p\} \in \mathbb{R}^n \end{aligned}$$

To say $\vec{b} \in span(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p)$ means that there is a solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots c_p\vec{v}_p = \vec{b}$.

3.2.1 Spans in $\mathbb{R}^3$

Now, we study the spans of three-vectors ($\mathbb{R}^3$).

Case 1: $\vec{v}_1$ is the 0 vector. $span(\vec{v}_1)$ is the origin.

Case 2: $\vec{v}_1 \neq 0$. $span(\vec{v}_1)$ is a straight line going through the origin.

Case 3: $\vec{v}_1 \neq 0$ and $\vec{v}_2 = c\vec{v}_1$. $span(\vec{v}_1, \vec{v}_2)$ is again a straight line going through the origin.

Case 4: $\vec{v}_1 \neq 0$ and $\vec{v}_2$ is not a multiple of $\vec{v}_1$. $span(\vec{v}_1, \vec{v}_2)$ is a plane going through the origin.

3.2.2 Matrices and Vectors

If we say A is an "$m \times n$" matrix, this means A has m rows and n columns. If A is an $m \times n$ matrix with columns $\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \end{bmatrix}$ and we have an $\mathbb{R}^m$ vector $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$, their product is defined as $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots x_m\vec{v}_m$. In other words, the ith entry of $A\vec{x}$ is the (ith row of A) $\cdot\vec{x}$.

For example, if we wish to write:

$$\begin{aligned} 8x_1 - x_2 &= 4 \\ 5x_1 + 4x_2 &= 1 \\ x_1 - 3x_2 &= 2 \end{aligned}$$

as a vector equation and as a matrix equation, we would do the following:

In vector equation form,

$$x_1 \begin{bmatrix} 8 & \\ 5 & \\ 1 & \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

And in matrix equation form,

$$\begin{bmatrix} 8 & -1 \\ 5 & 4 \\ 1 & -3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$$

So we have seen that if A is an $m \times n$ matrix with columns $\vec{a}_1, \vec{a}_2, \dots \vec{a}_n$ and $\vec{b}$ is an m-vector, then the solution set of $A\vec{x} = \vec{b}$ is the same as the solution set of $\sum_{i=1}^{n} x_i\vec{a}_i = \vec{b}$ is the same as the solution set of the augmented matrix $\left[\begin{array}{cccc|c} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n & \vec{b} \end{array} \right]$.

We have a theorem that if A is a $m \times n$ matrix, then the following are equivalent:

a. For all $\vec{b} \in \mathbb{R}^m$, the equation $A\vec{x} = \vec{b}$ has a solution.

b. Each $\vec{b} \in \mathbb{R}^m$ is a scalar 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 few rules about matrix multiplication: $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$, and $A(c\vec{u})) = cA\vec{u}$.

4 August 29 Lecture

In taking the product $A\vec{x}$, we have such a thing as an identity matrix: I_n:

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

This is an $n \times n$ matrix. The identity matrix has the property that $I\vec{x} = \vec{x}$.

4.1 Homogenous Linear Equations

We define a homogeneous linear equation as an equation of the form $A\vec{x} = 0$, where A is an $m \times n$ matrix, $\vec{x} \in \mathbb{R}^n$, and $\vec{0} \in \mathbb{R}^m$. We say that the homogeneous linear equation has a trivial solution $\vec{x} = 0$. Is there a non-trivial solution?

We can say that there will only be a non-trivial equation if there are free variables, because if every variable is non-free, all of them must equal 0.

For example, if we have the augmented matrix:

$$\left[\begin{array}{cccc|c} 1 & 3 & -3 & 7 & 0 \\ 0 & 1 & -4 & 5 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 9 & -8 & 0 \\ 0 & 1 & -4 & 5 & 0 \end{array}\right]$$

Then we have the solution:

$$\vec{x} = x_3 \begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}$$

4.2 Inhomogenous Linear Equations

These are equations of the form $A\vec{x} = \vec{b}$. If we suppose that $\vec{p}$ is a solution of $A\vec{x} = \vec{b}$ then any other solution is of form $\vec{p} + \vec{x}$, where $\vec{x}$ solves $A\vec{x} = 0$. If we wish to solve:

$$\left[\begin{array}{cccc|c} 1 & 3 & -3 & 7 & 8 \\ 0 & 1 & -4 & 5 & 2 \end{array}\right]$$

Say we know one solution is $x_1 = x_2 = x_3 = x_4 = 1$. We have solved the homogenous form of this equation above, so the general solution is:

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} -9 \\ 4 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}$$

4.3 Linear Dependence

We say a collection of vectors $\vec{v}_1, \vec{v}_2, \dots \vec{v}_p$ is linearly dependent if any one of them can be written as a linear combination of the others.

5 August 31 Lecture

Another definition for linear dependence is that the vectors $\vec{v}_1, \vec{v}_2, \dots \vec{v}_p$ is linearly dependent if there are numbers $x_1, x_2, \dots x_p$ not all zero, so that $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots x_p\vec{v}_p = 0$. We can show that these are equivalent definitions by saying:

$$\vec{v}_j = \sum_{i \neq j} c_i \vec{v}_i$$

This is the condition for our first definition of linear dependence. Now, if we simply set c_j to -1, we see that:

$$c_j\vec{v}_j + \sum_{i \neq j} c_i \vec{v}_i = 0$$

This matches our second definition.

5.1 Linear Independence

If $\vec{v}_1, \vec{v}_2, \dots \vec{v}_p$ are linearly independent, then whenever $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots x_p\vec{v}_p = 0$, all x_i must be equal to 0. Formally, if and only if the sole solution of $A\vec{x} = 0$ is $\vec{x} = 0$, then the columns of A are linearly independent.

If $p > n$ then any collection of p vectors in $\mathbb{R}^n$ is linearly dependent. Say:

$$A = \left[\begin{array}{c|c|c|c} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \end{array}\right]$$

We ask, does $A\vec{x} = 0$? Since A has at most n pivot points, we have *at least* $p - n$ free variables. Thus $A\vec{x} = 0$ must have a nontrivial solution, and the vectors are linearly dependent.

5.2 Matrix Transformations

Say A is a matrix. We ask, what is the matrix transformation corresponding to A? We can say this is a 'machine' that takes input $\vec{x}$ and outputs $A\vec{x}$. For example, say we have:

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

Now, we have:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Then, the matrix transformation of A is:

$$A\vec{x} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$$

If we graph the vector $\vec{x}$, and then graph the transformed vector $A\vec{x}$, in a sentence it would result in a reflection about the x-axis. If we had $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, then this would result in a matrix transformation of a reflection about the y-axis.

We say that if T is a transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, if takes a $\vec{x} \in \mathbb{R}^m$ and outputs $T(\vec{x}) \in \mathbb{R}^n$. Now, if we recall the rules for matrix-vector multiplication:

$$A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$$
$$A(c\vec{u}) = cA\vec{u}$$

We can say that a transformation T is linear if $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ and $T(c\vec{u}) = cT(\vec{u})$. As a side effect, if a transformation T is linear, then $T(0) = 0$ and $T(c\vec{u}+d\vec{v}) = cT(\vec{u}) + dT(\vec{v})$. So, any matrix transformation is linear.

Suppose T is linear, and:

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}2\\3\end{bmatrix}$$
$$T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}4\\5\end{bmatrix}$$
$$T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}6\\7\end{bmatrix}$$

What is $T\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)$?

$$\begin{aligned}
T\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right) &= x_1 T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) + x_2 T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) + x_3 T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right)\\
&= x_1\begin{bmatrix}2\\3\end{bmatrix} + x_2\begin{bmatrix}4\\5\end{bmatrix} + x_3\begin{bmatrix}6\\7\end{bmatrix}\\
&= \begin{bmatrix}2 & 4 & 6\\3 & 5 & 7\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}
\end{aligned}$$

6 September 5 Lecture

Recall that a matrix transformation $\vec{x} \to A\vec{x}$ is a linear transformation. We have another theorem that any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation, for some A.

To show this, we put:

$$\vec{e}_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, \vec{e}_2 = \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}, \vec{e}_n = \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$

For $\vec{x}$ in $\mathbb{R}^n$:

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \ldots + x_n\vec{e}_n$$
$$T(\vec{x}) = x_1T(\vec{e_1}) + x_2T(\vec{e}_2) + \ldots + x_nT(\vec{e}_n) = A\vec{x}$$

where

$$A = \left[\ T(\vec{e}_1)\ \middle|\ T(\vec{e}_2)\ \middle|\ \ldots\ \middle|\ T(\vec{e}_n)\ \right]$$

Now, say for example we have $T : \mathbb{R}^3 \to \mathbb{R}^2$. Say $T(\vec{e}_1) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, T(\vec{e}_2) = \begin{bmatrix} 4 \\ -7 \end{bmatrix}, T(\vec{e}_3) = \begin{bmatrix} -5 \\ 4 \end{bmatrix}$. The 'standard matrix' is thus:

$$\begin{bmatrix} 1 & 4 & -5 \\ 3 & -7 & 4 \end{bmatrix}$$

Since:

$$T(\vec{x}) = \begin{bmatrix} 1 & 4 & -5 \\ 3 & -7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Now, we define a transformation T to be "onto" if for every $\vec{b}$ in $\mathbb{R}^m$, there's some point $\vec{x}$ in $\mathbb{R}^n$ so that $T(\vec{x}) = \vec{b}$. We say T is one-to-one if for $\vec{b}$ in $\mathbb{R}^m$, there's at most one $\vec{x}$ so that $T(\vec{x}) = \vec{b}$. We prove that if the only solution to $T(\vec{x}) = 0$ is $\vec{x} = 0$, then T must be one to one. We do so by contradiction. Suppose that T is not one-to-one. Then there is a $\vec{u}$ and a $\vec{v}$ with $\vec{u} \neq \vec{v}$ so that $T(\vec{u}) = T(\vec{v}) = \vec{b}$. But now, $T(\vec{u} - \vec{v}) = 0$, and since $\vec{u} - \vec{v} \neq 0$, we have a solution for $T(\vec{x}) = 0$ that is not 0.

We have two theorems, with $T : \mathbb{R}^n \to \mathbb{R}^m$ which is linear and where A is the standard matrix of T. T is onto if and only if the columns of A span $\mathbb{R}^m$ (this follows from the definition of span and onto), and T is one-to-one if and only if the columns of A are linearly independent (this is because if they were *linearly dependent*, $T(\vec{x})$ would have nontrivial solutions, and we proved before that this means the system is not one-to-one).

6.1 Matrix Algebra

If we have a matrix A that is $m \times n$, then a_{ij} is the entry in the ith row and j column. If A and B are both $m \times n$ matrices, we define $(A + B)_{ij} = a_{ij} + b_{ij}$.

We define matrix multiplication, where A is an $m \times n$ matrix and B is a $n \times p$ matrix, we define AB (an $m \times p$ matrix). We require that $A(B\vec{x}) = (AB)\vec{x}$ for all $\vec{x} \in \mathbb{R}^p$, and that $A(BC) = (AB)C$, as long as the dimensions match up. Now, we define $(AB)_{ik}$ as the dot product of the $i$the row of A with the kth column of B.

7 September 7 Lecture

We note that if AB is defined, BA is not necessarily defined. Even if AB and BA are defined, it is possible that $AB \neq BA$.

We now discuss powers of matrices. If A is a $n \times n$ square matrix, $A^2 = AA$, $A^3 = AAA$, etc. We say that $A^0 = I_n$, where I is the identity matrix. A transpose of an $m \times n$ matrix A is defined as an $n \times m$ matrix: ${A^T}_{ij} = A_{ji}$. Now, we have several theorems:

- $(A^T)^T = A$
- $(rA)^T = rA^T$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Also, if $AB = AC$, it is not necessary that $B = C$. However, if we have $A(B\vec{x})$, we can necessarily say that $A(B\vec{x}) = (AB)\vec{x}$.

7.1 Matrix Inverses

Say we have A, a $n \times n$ matrix. We say A is invertible (or nonsingular) if there's an $n \times n$ matrix C so that $CA = AC = I_n$. We now show that C is unique. Say $BA = I$, $AB = I$. Now:

$$\begin{aligned} BA &= I \\ (BA)C &= IC \\ B(AC) &= C \\ BI &= C \\ B &= C \end{aligned}$$

Thus C must be unique.

We show how to find the inverse of a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ by introducing the determinant $\det(A) = ad - bc$. Then the inverse $A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. We note that if $A\vec{x} = \vec{b}$, then $\vec{x} = A^{-1}\vec{b}$. We show this by the following:

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A^{-1}(A\vec{x}) &= A^{-1}\vec{b} \\ (A^{-1}A)\vec{x} &= A^{-1}\vec{b} \\ \vec{x} &= A^{-1}\vec{b} \end{aligned}$$

We now introduce a few theorems involving inverses:

- If A is invertible, then so is A^{-1}, and $(A^{-1})^{-1} = A$. This is because if $AA^{-1} = I$, $A^{-1}A = I$, then $(A^{-1})^{-1} = A$.
- If A and B are invertible $n \times n$ matrices then so is AB, and $(AB)^{-1} = B^{-1}A^{-1}$. This is because

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= I \end{aligned}$$

Similarly, $(B^{-1})A^{-1}(AB) = I$.

- If A is invertible then so is A^T, and $(A^T)^{-1} = (A^{-1})^T$. This is because:

$$(A^{-1})^T = I, A^T(A^{-1})^T = I$$
$$\text{so, } (A^T)^{-1} = (A^{-1})^T$$

We can find inverses by making an $n \times 2n$ augmented matrix:

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

We do row reduction to get $\left[\begin{array}{c|c} I_n & A^{-1} \end{array}\right]$.

8 September 10 Lecture

We define a transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ to be invertible if there is some $S : \mathbb{R}^n \to \mathbb{R}^n$ so that $S(T(\vec{x})) = \vec{x}$ and $T(S(\vec{x})) = \vec{x}$ for any $\vec{x}$ in $\mathbb{R}^n$.

We say $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation with standard matrix A. Then T is invertible if and only if A is invertible. Then S has standard matrix A^{-1}. The proof is as follows: say T is invertible. We claim that T is one-to-one, because if $T(\vec{x}_1) = T(\vec{x}_2)$ and $\vec{x}_1 \neq \vec{x}_2$, then we have $S(T(\vec{x}_1)) = S(T(\vec{x}_2))$, but then this means $\vec{x}_1 = \vec{x}_2$. This is a contradiction, so T must be one to one. Also T is onto since given any $\vec{b}$ in $\mathbb{R}^n$, we can have $T(S(\vec{b}) = \vec{b}$ and $S(\vec{b})$ is well defined for all $\mathbb{R}^n$. So T is one-to-one and onto, thus A is invertible.

8.1 More on Determinants

We can find the determinant of a 3×3 matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The determinant is $|A| = a_{11} \cdot \det\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right) - a_{12} \cdot \det\left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}\right) + a_{13} \cdot \det\left(\begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{32} \end{bmatrix}\right)$.

More generally, given an $n \times n$ matrix A, we say A_{ij} is the $(n-1) \times (n-1)$ matrix where we remove the ith row and jth column of A. Then, the determinant of A is:

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \ldots + (-1)^{n+1}a_{1n}\det(A_{1n})$$

We define the cofactor for a matrix entry $C_{ij} = (-1)^{i+j}\det(A_{ij})$. By definition:

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}$$

However, now we can actually compute the determinant of A by "cofactor expansion" along any row or column. Before we were doing is along the top row, and now we can do it along any row or column!

9 September 12 Lecture

If A is a 'triangular' matrix, then $\det(A)$ is the product of its diagonalentries. Determinants have a few properties:

1. $\det(A^T) = \det(A)$. This is because we can calculate determinant along rows or columns.
2. If we switch two rows, $\det(A)$ changes by a sign.
3. If we multiply a row (or column) by a constant c then the determinant gets multiplied by c.
4. If we add a multiple of one row to another, this does not change the determinant at all.

We have a theorem that A is invertible if and only if A $\neq 0$. We prove this by noting that whether or not A is invertible is unchanged by row operations. Thus we can assume A is upper triangular. A is invertible if and only if A has n pivot positions, meaning all the diagonal terms are non-zero. If any entry on a diagonal is 0, then it is non-invertible and the determinant is 0.

A few more facts about determinants:

1. $\det(AB) = \det(A)\det(B)$
2. $\det(A^{-1}) = \frac{1}{\det(A)}$ if A is invertible.

We have a method of computing inverses by making an augmented matrix with the right side being the identity matrix and the left side being A. Another way is to make a cofactor matrix:

$$\begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} \\ C_{21} & C_{22} & \dots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} \end{bmatrix}$$

The adjugate matrix $\mathrm{adj}(A) = C^T$. Now, we say $A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}$

10 September 14 Lecture

10.1 The Geometric Meaning of Determinants

We have a notion of a parallelogram generated by two vectors in $\vec{u}$, $\vec{v}$ in $\mathbb{R}^2$ – it has vertices $\vec{0}$, $\vec{u}$, $\vec{v}$, and $\vec{u}+\vec{v}$. A parallelpiped is generated by vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ in $\mathbb{R}^3$. We say that if A is a 2×2 matrix then the area of the parallelogram generated by its columns is equal to $\det(A)$. The same goes for A if it is a 3×3 matrix – the determinant is the volume of the parallelpiped generated by its columns.

We prove this as follows. We say $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then $\vec{u} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$, and $\vec{v} = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$. Now, we know that the area between two vectors is $\|\vec{u}\|\|\vec{v}\|\sin\theta = \|\vec{u}\times\vec{v}\|$. Now, $\vec{u}\times\vec{v} = \begin{bmatrix} 0 \\ 0 \\ a_{11}a_{22} - a12a_{21} \end{bmatrix}$. Now, $|\vec{u}\times\vec{v}| = a_{11}a_{22} - a12a_{21} = \det(A)$.

Alternatively, we may define a determinant as an 'area distortion'. We say A is a 2×2 matrix so that $T : \mathbb{R}^2 \to \mathbb{R}^2$, then when we apply this transformation to some object of area α, we will end up with an object of area $\det(A)\alpha$.

For example, if we have a unit circle and apply the linear transformation $T(\vec{x}) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \vec{x}$, then the result is an ellipse with axes a and b. The area is now $ab\pi$, since the determinant of the matrix is ab. The same is true for an ellipsoid of axes a, b, and c. The area is $\frac{4}{3}\pi abc$.

10.2 Cramer's Rule

We have Cramer's rule to solve $A\vec{x} = \vec{b}$, when A is an invertible $n \times n$ matrix. If x_i represents the ith entry of the solution for $\vec{x}$, its value is $\frac{\det(B)}{\det(A)}$, where B is what we get when we replace the ith column of A with $\vec{b}$.

11 September 17 Lecture

There are some examples of vector spaces:

- $\mathbb{R}^n$
- $\mathbb{P}_n$, which is the set of polynomials of degree at most n.
- $\mathbb{R}^\infty$, which is the set of all sequences $a_1, a_2, a_3, \ldots$, where a_i is a real number for all $i \geq 1$
- $\mathbb{S}$, which is the set of all bi-infinite sequences $\ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots$
- $\mathbb{V}$, which is the set of all functions of a variable x: $(f_1 + f_2)(x) = f_1(x) + f_2(x)$, and $(cf)(x) = cf(x)$.
- $\mathbb{C}(\mathbb{R})$ is the set of all continuous functions of one variable

We define a subspace of $\mathbb{V}$, where $\mathbb{V}$ is a vector space, as a subset $\mathbb{H}$ of $\mathbb{V}$ so that:

1. 0 is in $\mathbb{H}$
2. If u and v are in $\mathbb{H}$, then $u + v$ is in $\mathbb{H}$
3. If u is in $\mathbb{H}$ and $c \in \mathbb{R}$, then cu is in $\mathbb{H}$

For example, if we have vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p \in \mathbb{V}$, then $\text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p)$ is a subspace of $\mathbb{V}$.

We now define a linear transformation T from a vector space $\mathbb{V}$ to a vector space $\mathbb{W}$ as a rule that assigns to each $\vec{v} \in \mathbb{V}$ some $T(\vec{v})$ in $\mathbb{W}$, so that $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ and $T(c\vec{u}) = cT(\vec{u})$. Two examples are the matrix transformation $\vec{x} \to A\vec{x}$ and the derivative operator $\frac{\mathrm{d}}{\mathrm{d}x}$.

We define a kernel (or nullspace) of T as the set of all $u\vec{V}$ so that $T(\vec{u}) = 0_{\mathbb{W}}$. Now, the range of T is the set of all $T(\vec{x})$ where $\vec{x} \in \mathbb{V}$. We can show that the kernel is a subspace of $\mathbb{V}$ and range is a subspace of $\mathbb{W}$.

12 September 19 Lecture

We introduce notation $\vec{x} \to A\vec{x}$ as a matrix transformation, where $\text{Nul}(A)$ represents the kernel and $\text{Col}(A)$ represents range.

A linear transformation T is one-to-one if and only if the kernel of T is only the 0 vector. It is onto if and only if the range is spans all of $\mathbb{W}$.

12.1 Bases

We define the basis of a vector space $\mathbb{V}$ as a set $\mathbb{B} = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_p\right\}$ of vectors so that $\mathbb{B}$ is linearly independent and $\text{span}(\mathbb{B}) = \mathbb{V}$.

Let $\mathbb{S} = \{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p\}$ be a subset of $\mathbb{V}$. We say $\mathbb{H} = \text{span}(\vec{v}_1, \vec{v}_2, \ldots \vec{v}_p)$. First, we can say that if one of the vectors in S, say $\vec{v}_k$, is a linear combination of the others, then we remove it, then what's left still spans $\mathbb{H}$. If $\mathbb{H} \neq \{0\}$, then some subset of $\mathbb{S}$ is a basis of $\mathbb{H}$ (we do this by removing vectors from $\mathbb{S}$ until it is linearly independent – then its span is $\mathbb{H}$ since we don't lose anything by removing linearly dependent vectors).

13 September 21 Lecture

A basis $\mathbb{B} = \left\{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_p\right\}$ consists of linearly independent vectors that span $\mathbb{V}$. It is not unique for $\mathbb{V}$, but every $\mathbb{V}$ has a basis. With reference to $\vec{x} \to A\vec{x}$, a kernel is $\text{Nul}(A) = \{\vec{x} \text{ so that } A\vec{x} = 0\}$ and the image is $\text{Col}(A)$ = span of column vectors of A. To get the basis of the column space $\text{Col}(A)$, we choose the columns of A that correspond to pivot positions (since non-pivot positions indicate free variables, and thus their corresponding columns are redundant).

We have a theorem that if $\mathbb{B} = \left\{\vec{v}_1, \ldots, \vec{b}_p\right\}$ of $\mathbb{V}$, then every $\vec{x}$ in $\mathbb{V}$ can be written **uniquely** as $\vec{x} = \sum_{i=1}^{p} c_i \vec{b}_i$ for scalars c_i. We show this as follows. Because $\mathbb{B}$ spans V, we can write $\vec{x} = \sum_{i=1}^{p} c_i \vec{b}_i$. However, we must show it is unique! Suppose $\vec{x} = \sum_{i=1}^{p} c_i' \vec{b}_i$ is another way to write $\vec{x}$. We subtract: $\vec{0} = \sum_{i=1}^{p} (c_i - c_i')\vec{b}_i$. However, by linearly independence of $\mathbb{B}$, all $c_i - c_i'$ **must be** 0 thus $c_i = c_i'$ and we have only way way to write $\vec{x}$. We define $[\vec{x}]_{\mathbb{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$, which are the coordinates of $\vec{x}$ relative to $\mathbb{B}$.

Now, say for example $\mathbb{V} = \mathbb{R}^n$, we say $\mathbb{B} = \left\{\vec{b}_1, \ldots, \vec{b}_p\right\}$. Given $\vec{x} \in \mathbb{R}^n$, we can write $\vec{x} = \sum_{i=1}^{p} c_i \vec{b}_i$. Thus $\vec{x} = \left[\begin{array}{c|c|c|c} \vec{b}_1 & \vec{b}_2 & \ldots & \vec{b}_p \end{array}\right] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = P_{\mathbb{B}}[\vec{x}]_{\mathbb{B}}$. Now, we can also say $\vec{x} = P_{\mathbb{B}}[\vec{x}]_{\mathbb{B}} \to [\vec{x}]_{\mathbb{B}} = P_{\mathbb{B}}^{-1}\vec{x}$. Thus We call $P_{\mathbb{B}}$ the 'change of coordinate matrix' from $\mathbb{B}$ to the standard basis.

We define an isomorphism between two vector spaces as a one-to-one and onto linear mapping.

14 September 24 Lecture

We define the column and row space of a matrix A to be $\text{Col}(A)$ and $\text{Row}(A)$, which is the span of the column and row vectors of A, respectively. The dimension of both of these is equal to the number of pivot positions in an echelon form of A.

We have a 'basis' theorem: if we suppose that $\dim V = p$, then any linearly independent set of p vectors is a basis of V, and any set of p elements that spans V is also a basis.

We define the rank of a matrix to be the dimension of $\text{Col}A$. We have a rank theorem:

- If A is an $m \times n$ matrix, then $n = \text{Rank}(A) + \dim(\text{Nul}(A))$.
- $\dim(\text{Row}(A)) = \text{Rank}(A)$.

15 September 26 Lecture

The rank theorem holds that if A is an $m \times n$ matrix, then $n = \text{Rank}(A) + \dim(\text{Nul}(A))$ and $\dim(\text{Row}(A)) = \text{rank}(A)$.

For example, suppose that A is a 5×6 matrix that has 4 pivot columns. What is $\dim(\text{Nul}(A))$? Is $\text{Col}(A) = \mathbb{R}^4$? The nullity of A is 2, the dimension of the column space of A is 4. However, $\text{Col}(A)$ is a subspace of $\mathbb{R}^5$ – however, it is isomorphic to $\mathbb{R}^4$.

15.1 Change of Coordinates

Say $\vec{x}$ is a vector in a vector space $\mathbb{V}$. Suppose that $\mathbb{B} = \left\{\vec{b}_1, \ldots, \vec{b}_n\right\}$ and $\mathbb{G} = \{\vec{c}_1, \ldots, \vec{c}_n\}$ are bases of $\mathbb{V}$. $[\vec{x}]_{\mathbb{B}}$ is a column bector – we can write $\vec{x} = x_1\vec{b}_1 + \ldots + x_n\vec{b}_n$, then $[\vec{x}]_{\mathbb{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Now, there is some $n \times n$ matrix $P_{\mathbb{G}\leftarrow\mathbb{B}}$ so for all $\vec{x} \in \mathbb{V}$, $[\vec{x}]_{\mathbb{G}} = (P_{\mathbb{G}\leftarrow\mathbb{B}})[\vec{x}]_{\mathbb{B}}$. We note that we can find this matrix $P_{\mathbb{G}\leftarrow\mathbb{B}}$ since its first column is $[\vec{b}_1]_{\mathbb{G}}$, its second column is $[\vec{b}_2]_{\mathbb{G}}$, and so on. We can also do this by doing row operations on $\left[\begin{array}{ccc|ccc} \vec{c}_1 & \ldots & \vec{c}_n & \vec{b}_1 & \ldots & \vec{b}_n \end{array}\right]$ into $\left[\begin{array}{c|c} I_n & P_{\mathbb{G}\leftarrow\mathbb{B}]} \end{array}\right]$

16 October 1 Lecture

We define an eigenvector of a square matrix A to be a nonzero vector $\vec{x}$ so that $A\vec{x} = \lambda\vec{x}$ for some scalar λ. A scalar λ is an eigenvalue of A if there is a non-zero solution of $A\vec{x} = \lambda\vec{x}$.

We have a theorem that for any eigenvalue λ, the corresponding eigenvectors, along with $\vec{0}$, form a subspace. We show this by noting if $A\vec{x} = \lambda\vec{x}$, then $\vec{x}$ is in the nullspace of $A - \lambda I$. A nullspace is always a subspace.

We note that the eigenvalues of a diagonal matrix are the values along its diagonal. This is because if the values along the diagonal are $d_1, d_2, \ldots, d_n$, then the values along the diagonal of $A - \lambda I$ are $d_1 - \lambda, d_2 - \lambda, \ldots, d_n - \lambda$, and the determinant is then $(d_1 - \lambda)(d_2 - \lambda)\ldots(d_n - \lambda)$, and now λ can be any of $d_1, d_2, \ldots, d_n$ (since we set determinant to be 0). It is possible to have 'double eigenvalues' (identical eigenvalues for different eigenvectors).

16.1 Similarity

We define square matrices A and B to be similar if there's an invertible matrix P so that $P^{-1}AP = B$. We have a theorem that if A and B are similar, they have the same eigenvalues. We do this as follows:

$$B - \lambda I = P^{-1}A^{P} - \lambda I$$

Since $\lambda I = P^{-1}P\lambda I = P\lambda I P^{-1}$

$$\begin{aligned}
&= P^{-1}AP - P^{-1}\lambda I P \\
&= P^{-1}(A - \lambda I)P \\
\det(B - \lambda I) &= \det(P^{-1}(A - \lambda I)P) \\
&= \det(P^{-1})\det(A - \lambda I)\det(P) \\
&= \det(P^{-1}P)\det(A - \lambda I) \\
&= \det(A - \lambda I)
\end{aligned}$$

Since their determinant is the same, they must have identical eigenvalues.

17 October 3 Lecture

We have a theorem that if $\{v_1, \ldots, \vec{v}_p\}$ are eigenvectors corresponding to distinct eigenvalues, then they are linearly independent.

We have the Fibonnaci sequence 0, 1, 1, 2, 3, 5, ... If $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\vec{x}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and so on, then we can say:

$$\begin{aligned}
\vec{x}_{n+1} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \vec{x}_n \\
\vec{x}_n &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \vec{x}_0
\end{aligned}$$

The eigenvalues for this matrix are $\lambda = \frac{1+\pm\sqrt{5}}{2}$, and corresponding eigenvectors $\vec{v}_1, \vec{v}_2$. Now:

$$\begin{aligned}
A^n\vec{x}_0 &= A^n(c_1\vec{v}_1 + c_2\vec{v}_2) \\
&= c_1A^n\vec{v}_1 + c_2A^n\vec{v}_2 \\
&= c_1\left(\frac{1+\sqrt{5}}{2}\right)^n \vec{v}_1 + c_2\left(\frac{1-\sqrt{5}}{2}\right)^n \vec{v}_2
\end{aligned}$$

In the end, we get

$$u_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

We define a square matrix A to be diagonalizable if $A = PDP^{-1}$ for some diagonal matrix D. These are useful because powers of a diagonalizable matrix are easy:

$$\begin{aligned}
A^2 &= (PDP^{-1})(PDP^{-1}) \\
&= PD(P^{-1}P)DP^{-1} \\
&= PDIDP^{-1} \\
&= PD^2P^{-1} \\
A^k &= PD^kP^{-1}
\end{aligned}$$

We have a thorem that A is diagonalizable if and only if A has n linearly independent eigenvectors. Then $A = PDP^{-1}$ where the columns of P are the eigenvectors and the diagonal entries of D are eigenvalues.

18 October 8 Lecture

When we are searching for eigenvalues, we may sometimes find imaginary eigenvalues and eigenvectors. There will always be n eigenvalues if we allow them to be complex, since the characteristic polynomial is always of nth degree. We have a special case, where $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, with $b \neq 0$. The determinant of $A - \lambda I = (a-\lambda)^2 + b^2$. Setting this to 0:

$$\begin{aligned}
(a-\lambda)^2 &= -b^2 \\
\lambda &= a \pm ib
\end{aligned}$$

Now, $A - \lambda I = \begin{bmatrix} ib & -b \\ b & ib \end{bmatrix}$. We row reduce this:

$$\left[\begin{array}{cc|c} ib & -b & 0 \\ b & ib & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & i & 0 \\ 1 & i & 0 \end{array}\right]$$

Thus we have $x_1 = -ix_2$, or $\vec{x} = \left\{\alpha \begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$, $\alpha \in \mathbb{C}$. We can now write $A = \begin{bmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{bmatrix}$, where $r = \sqrt{a^2+b^2}$, and $\theta = \arctan\frac{b}{a}$. Thus, a is an expansion by r and a rotation by θ.

Say we have A, a real $n \times n$ matrix. Say λ is a complex eigenvalue with complex eigenvector x. Now, $Ax = \lambda x$, but $A\overline{x} = \lambda\overline{x}$. We have a theorem that eigenvalues always come in conjugate pairs.

Say we have a two by two matrix that has no real eigenvalues (and thus 2 complex eigenvalues that are conjugate pairs). Say $\lambda = a \pm ib$, and $\vec{v}$ is an eigenvector corresponding to $\lambda = a - ib$. We let $P = \left[\begin{array}{c|c} \mathrm{Re}(\vec{v}) & \mathrm{Im}(\vec{v}) \end{array}\right]$. Then $A = PCP^{-1}$, where $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, so A is a combination of rotation and expansion.

Now we define a dot product: if we have $\vec{u} = \begin{bmatrix} u_1 \\ \dots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix}$, and $\vec{u}\cdot\vec{v} = u_1v_1 + \ldots + u_nv_n$. Now, we can say the length of a vector $\vec{u}$ as $\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}}$. Finally, we define the distance between two vectors as $\mathrm{dist}(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|)$. We define $\vec{u}$ and $\vec{v}$ to be orthogonal if $\vec{u}\cdot\vec{v} = 0$.

19 October 10 Lecture

We have a Pythagorean theorem in n dimensions: $\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$ if and only if $\vec{x}\cdot\vec{y} = 0$. We prove this as follows:

$$\begin{aligned}
\|\vec{x}+\vec{y}\|^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2 &= (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) - \vec{x}\cdot\vec{x} - \vec{y}\cdot\vec{y} \\
&= (\vec{x}\cdot\vec{x} + \vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{x} + \vec{y}\cdot\vec{y}) - \vec{x}\cdot\vec{x} - \vec{y}\cdot\vec{y} \\
&= \vec{x}\cdot\vec{y} - \vec{y}\cdot\vec{x} \\
&= 0
\end{aligned}$$

19.1 Orthogonality

Say W is a subspace of $\mathbb{R}^n$, then $W^\perp = \{\vec{z}\in\mathbb{R}^n \text{ so } \vec{z}\cdot\vec{w} = 0 \text{ for all } \vec{w}\in W\}$. We now show $W^\perp$ is a subspace of $\mathbb{R}^n$. First, we show it is closed under addition. Say $\vec{z}_1 \in W^\perp$ and $\vec{z}_2 \in W^\perp$, then we have $\vec{z}_1\cdot\vec{w} = 0$ and $\vec{z}_2\cdot\vec{w} = 0$ and thus $(\vec{z}_1+\vec{z}_2)\cdot\vec{w} = \vec{z}_1\cdot\vec{w} + \vec{z}_2\cdot\vec{w} = 0$. Similarly for scalar multiplication, if $\vec{z}\cdot\vec{w} = 0$, then $c\vec{z}\cdot\vec{w} = 0$.

The Cauchy-Schwartz inequality for any $\vec{u},\vec{v}\in\mathbb{R}^n$ is:

$$\|\vec{u}\cdot\vec{v}\| \leq \|\vec{u}\|\ \|\vec{v}\|$$

We show this by:

$$\begin{aligned}
0 &\leq \|a\vec{u} - b\vec{v}\| \\
&\leq a^2\|\vec{u}\|^2 - 2ab\vec{u}\cdot\vec{v} + b^2\|\vec{v}\|^2 \\
\vec{u}\cdot\vec{v} &\leq \frac{a^2\|\vec{u}\|^2 + b^2\|\vec{v}\|^2}{2ab}
\end{aligned}$$

Since our inequality was true for all a and b, we can take $a = \|\vec{v}\|$ and $b = \|\vec{u}\|$:

$$\begin{aligned}
\vec{u}\cdot\vec{v} &\leq \frac{\|\vec{u}\|^2\|\vec{v}\|^2 + \|\vec{u}\|^2\|\vec{v}\|^2}{2\|\vec{u}\|\ \|\vec{v}\|} \\
&\leq \|\vec{u}\|\ \|\vec{v}\|
\end{aligned}$$

Since this works for all $\vec{u}$ and $\vec{v}$, we sub in $\vec{u} = -\vec{u}$ and get:

$$\begin{aligned}
-\vec{u}\cdot\vec{v} &\leq \|\vec{u}\|\ \|\vec{v}\| \\
\vec{u}\cdot\vec{v} &\geq -\|\vec{u}\|\ \|\vec{v}\| \\
-\|\vec{u}\|\ \|\vec{v}\| &\leq \vec{u}\cdot\vec{v} \leq \|\vec{u}\|\ \|\vec{v}\| \\
\|\vec{u}\cdot\vec{v}\| &\leq \|\vec{u}\|\ \|\vec{v}\|
\end{aligned}$$

A set of vectors $\{\vec{u}_1,\ldots,\vec{u}_p\}$ in $\mathbb{R}^n$ is orthogonal if $\vec{u}_j\cdot\vec{u}_j = 0$ if $i \neq j$. It is orthonormal if **in addition**, $\|\vec{u}_i\| = 1$ for all i. We can always turn an orthogonal set into an orthonormal by 'normalizing' the set by dividing each vector by their lengths.

If W is a subspace of $\mathbb{R}^n$, then an orthogonal basis of W is a basis which is orthogonal (likewise for orthonormal). If we have $\{\vec{u}_1, \ldots, \vec{u}_p\}$ that is an orthogonal basis for W, then for any $\vec{y} \in W$, we can write $\vec{y} = \sum_{i=1}^{p} c_i \vec{u}_i$ where $c_i = \frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2}$. We show this by writing:

$$\vec{y} = \sum_{i=1}^{p} c_i \vec{u}_i$$

Since $\vec{u}_i \cdot \vec{u}_1 = 0$ for all $i \neq 1$:

$$\begin{aligned} \vec{u}_1 \cdot \vec{y} &= \sum_{i=1}^{p} c_i \vec{u}_1 \cdot \vec{u}_i = c_1 \vec{u}_1 \cdot \vec{u}_1 \\ c_1 &= \frac{\vec{u}_1 \cdot \vec{y}}{\|\vec{u}_1\|^2} \end{aligned}$$

If we have an orthonormal basis, it simplifies even further to $c_i = \vec{u}_i \cdot \vec{y}$.

We have a rotation matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Before, we thought of this as a transformation $\vec{x} \to A\vec{x}$ and the inverse of $A^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = A^T$. This is interesting – this gives a general notion of an $n \times n$ matrix A as orthogonal if it is invertible and $A^{-1} = A^T$. Two examples are the 2×2 rotation matrix, and another is a reflection matrix such as $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, yet another is a composition of two rotation matrices:

$$\begin{bmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & \cos\theta_2 & -\sin\theta_2 \\ 0 & 0 & \sin\theta_2 & \cos\theta_2 \end{bmatrix}$$

20 October 12 Lecture

We define a square matrix A as orthogonal if it is invertible and $A^{-1} = A^T$. Now, we have a theorem that if A is orthogonal, then for all $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$$

We prove this by noting $\vec{x}^T \vec{y} = \vec{x} \cdot \vec{y}$, then:

$$\begin{aligned} (A\vec{x}) \cdot (A\vec{y}) &= (A\vec{x})^T A\vec{y} \\ &= \vec{x}^T A^T A\vec{y} \\ &= \vec{x}^T A^{-1} A\vec{y} \\ &= \vec{x}^T \vec{y} \\ &= \vec{x} \cdot \vec{y} \end{aligned}$$

Thus, if A is orthogonal, then for all $\vec{x} \in \mathbb{R}^n$, $\|A\vec{x}\| = \|\vec{x}\|$. This follows from letting $\vec{y} = \vec{x}$ in the above proof:

$$\begin{aligned} (A\vec{x}) \cdot (A\vec{x}) &= \vec{x} \cdot \vec{x} \\ \|A\vec{x}\|^2 &= \|\vec{x}\|^2 \\ \|A\vec{x}\| &= \|\vec{x}\| \end{aligned}$$

So if A is an orthogonal matrix transformation, $\vec{x} \to A\vec{x}$ preserves length. The converse of this is also true.

We now make the claim that if A is orthogonal, then its columns form an orthonormal basis of $\mathbb{R}^n$. Since we know A is invertible, if $e_1, \ldots, e_n$ are the unit vectors that form the basis for $\mathbb{R}^n$, then the columns of A form a basis:

$$A = \left[\begin{array}{c|c|c} A\vec{e}_1 & \ldots & A\vec{e}_n \end{array} \right]$$

Now, $(A\vec{e}_i) \cdot (A\vec{e}_j) = \vec{e}_i \cdot \vec{e}_j$. This dot product is 1 if $i = j$ and 0 otherwise, and thus $A\vec{e}_1, \ldots, A\vec{e}_n$ is orthonormal.

20.1 Orthonormal Projection

Say we wish to find the component of $\vec{y}$ in the $\vec{u}$ direction. Say $\vec{u} \neq \vec{0}$ and $\vec{y} \in \mathbb{R}^n$. We want to write $\vec{y} = \hat{y} + \hat{z}$, where $\hat{y}$ is a multiple of $\vec{u}$ and $\vec{z}$ is orthogonal to $\vec{u}$. We want $\hat{y} = \alpha\vec{u}$ for some α, so we say $\vec{y} = \alpha\vec{u} + \hat{z}$:

$$\begin{aligned} \vec{y} &= \alpha\vec{u} + \hat{z} \\ \vec{u} \cdot \vec{y} &= \alpha\vec{u} \cdot \vec{u} + \vec{u} \cdot \hat{z} \end{aligned}$$

Since $\hat{z}$ is orthogonal to $\vec{u}$, $\vec{u} \cdot \hat{z} = 0$.

$$\begin{aligned} \vec{u} \cdot \vec{y} &= \alpha\vec{u} \cdot \vec{u} \\ \alpha &= \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \\ \hat{y} &= \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}}\vec{u} \\ \hat{z} &= \vec{y} - \hat{y} \end{aligned}$$

Note that if $\vec{u} \to c\vec{u}$, $\hat{y}$ does not change. So now we have notation that $\hat{y} = \text{proj}_L(\vec{y})$ if L is the line containing $c\vec{u}$. We note that $\hat{y}$ is the point on L that is closest to the endpoint of $\vec{y}$.

We now wish to extend to the general case of a vector and a p-dimensional subspace. We do this by $\{\vec{u}_1, \ldots, \vec{u}_p\}$. This is simply:

$$\hat{y} = \sum_{i=1}^{p} \frac{\vec{y} \cdot \vec{u}_i}{\vec{u}_i \cdot \vec{u}_i}\vec{u}_i$$

21 October 15 Lecture

Say W is a subspace of $\mathbb{R}^n$. How do we get an orthogonal basis of W? We say $\{\vec{x}_1, \ldots, \vec{x}_p\}$ is a basis of W, and we wish to modify this to get an orthogonal basis $\{\vec{v}_1, \ldots, \vec{v}_p\}$. We do this by taking $\vec{v}_1 = \vec{x}_1$, and then taking $\vec{v}_2 = \vec{x}_2 - (\text{part of } \vec{x}_2 \text{ in } \vec{v}_1 \text{ direction}) = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1$. Now we take $\vec{x}_3 = \vec{x}_3 - (\text{part of } \vec{x}_3 \text{ in span}(\vec{v}_1, \vec{v}_2)) = \vec{x}_3 - \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2$. Now we have $\{\vec{v}_1, \dots, \vec{v}_p\}$.

22 October 17 Lecture

Say we have some points on a plane (x_i, y_i), with $i = 1, \dots, m$, and we want to fit this data with a straight line ($y = \beta_0 + \beta_1 x$). We define square error as the sum of $\sum_{i=1}^{m}(y_i - \beta_0 - \beta_1 x_i)^2$. The goal is to find β_0 and β_1 to minimize square error, and this will be called the least squared error. We can apply matrices to this. If we take:

$$\left\| \begin{matrix} y_1 - \beta_0 - \beta_1 x_1 \\ \vdots \\ y_m - \beta_0 - \beta_1 x_m \end{matrix} \right\|^2 = \left\| \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} - \beta_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - \beta_1 \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \right\|^2$$
$$= \left\| \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \right\|^2$$

This length squared gives us the squared error. Our goal is to find the column vector $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$ that minimizes the error. In general, our problem is to find $\vec{x}$ that minimizes $\left\|\vec{b} - A\vec{x}\right\|^2$, with A is an $m \times n$ matrix $\vec{x} \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^m$. We want to find $\hat{b} \in \text{Col}(A)$ so that $\left\|\vec{b} - \hat{b}\right\|^2$ is minimized. We can do this by projection (since if we want $\vec{b}$ to be closest to a point on a plane, we project $\vec{b}$ onto it)!

Thus $\hat{b} = \text{proj}_{\text{Col}(A)}\vec{b}$, and thus $\vec{b} - \hat{b} \in \text{Col}(A)^\perp$. So for every vector $A\vec{z} \in \text{Col}(A)$, we want $(A\vec{z}) \cdot (\vec{b} - \hat{b}) = 0$, since $\vec{b} - \hat{b}$ must be perpendicular to the column space of A. Thus:

$$\begin{aligned} (A\vec{z}) \cdot (\vec{b} - \hat{b}) &= 0 \\ (A\vec{z})^T(\vec{b} - \hat{b}) &= 0 \\ \vec{z}^T A^T(\vec{b} - \hat{b}) &= 0 \\ \vec{z} \cdot A^T(\vec{b} - \hat{b}) &= 0 \end{aligned}$$

Since this must be true for all $\vec{z} \in \mathbb{R}^n$:

$$\begin{aligned} A^T(\vec{b} - \hat{b}) &= 0 \\ A^T\vec{b} &= A^T\hat{b} \end{aligned}$$

Since $\hat{b} \in \text{Col}(A)$, we have $\hat{b} = A\vec{x}$:

$$A^T\vec{b} = A^T A\vec{x}$$

If A^TA is invertible, the answer is $\vec{x} = (A^TA)^{-1}A^T\vec{b}$.

If $A^TA$is not invertible, then the minimizer $\vec{x}$ is not unique.

23 October 19 Lecture

We define an inner product on V as an operation that assigns to any $\vec{x}, \vec{y} \in V$ as a number $\langle a, b\rangle$ so that:

- $\langle \vec{x}, \vec{y}\rangle = \langle \vec{y}, \vec{z}\rangle$
- $\langle \vec{x} + \vec{y}, \vec{z}\rangle = \langle \vec{x}, \vec{x}\rangle + \langle \vec{y}, \vec{z}\rangle$
- $\langle c\vec{x}, \vec{y}\rangle = c\langle \vec{x}, \vec{\rangle} y$
- $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0$ if and only if $\vec{x} = 0$

If we have $V = \mathbb{R}^n$, then $\langle x, y\rangle = \vec{x} \cdot \vec{y}$. What if we have functions instead? For example, $V = C[a, b]$ (continuously valued functions on the interval $a, b)$, and $f(t), g(t) \in V$ then we have $\langle f, g\rangle = \int_a^b f(t)g(t)\,\mathrm{d}t$. We can apply this definition to the four conditions for an inner product above, to see that it is indeed an inner product.

Everything we can do we a dot product we can do with inner product. For example, we define $\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x}\rangle}$. We define $\vec{x}$ and $\vec{y}$ as orthogonal if $\langle \vec{x}, \vec{y}\rangle = 0$. Also, the Cauchy-Shwartz inequality: $\|\langle \vec{x}, \vec{y}\rangle\| \leq \|\vec{x}\|\|\vec{y}\|$. We also have the triangle inequality, which says $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$. We prove this in an inner product space as follows:

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y}\rangle \\ &= \langle \vec{x}, \vec{x}\rangle + 2\langle \vec{x}, \vec{y}\rangle + \langle \vec{y}, \vec{y}\rangle \end{aligned}$$

By Cauchy-Shwartz inequality,

$$\begin{aligned} &\geq \|\vec{x}\|^2 + 2\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2 \\ &= (\|\vec{x}\| + \|\vec{y}\|)^2 \end{aligned}$$

We can also do Gram-Schmidt: we say $\{\vec{x}_1, \ldots, \vec{x}_p\}$ is a basis of V. What is the orthogonal basis $\{\vec{v}_1, \ldots, \vec{v}_p\}$? Put $\vec{v}_1 = \vec{x}_1$:

$$\begin{aligned} \vec{v}_2 &= \vec{x}_2 - \frac{\langle \vec{x}_2, \vec{v}_1\rangle}{\langle \vec{v}_1, \vec{v}_1\rangle}\vec{v}_1 \\ \vec{v}_3 &= \vec{x}_3 - \frac{\langle \vec{x}_1, \vec{v}_1\rangle}{\langle \vec{v}_2, \vec{v}_1\rangle}\vec{v}_1 - \frac{\langle \vec{x}_3, \vec{v}_2\rangle}{\langle \vec{v}_2, \vec{v}_2\rangle}\vec{v}_2 \end{aligned}$$

For example, say $V = C[-1,1]$. Then $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$. Finding an orthogonal basis for the subspace spanned by $\left\{1, t, t^2\right\}$:

$$\begin{aligned}
\vec{v}_1 &= \vec{x}_1 = 1 \\
\vec{v}_2 &= t - \frac{\langle t, 1\rangle}{\langle 1, 1\rangle}1 \\
&= t - \frac{\int_{-1}^{1} t\,\mathrm{d}t}{\int_{-1}^{1} 1\,\mathrm{d}t} \\
&= t \\
\vec{v}_3 &= t^2 - \frac{\langle t^2, 1\rangle}{\langle 1, 1\rangle}1 - \frac{\langle t^2, t\rangle}{\langle t, t\rangle}t \\
&= t^2 - \frac{\int_{-1}^{1} t^2}{\int_{-1}^{1} 1\,\mathrm{d}t}1 - \frac{\int_{-1}^{1} t^3\,\mathrm{d}t}{\int_{-1}^{1} t^2\,\mathrm{d}t} \\
&= t^2 - \frac{1}{3}
\end{aligned}$$

Thus the Legendre polynomials are:

$$1, t, \frac{1}{2}(3t^2 - 1), \ldots$$

24 October 22 Lecture

A symmetric matrix is a square matrix A such that $A^T = A$. We define a square matrix as orthogonally diagonalizable if it has an orthogonal basis of real eigenvectors (i.e. $A = PDP^{-1}$). Since $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal basis, P is orthogonal ($P^{-1} = P^T$). We have a theorem that A is symmetric if and only if A is orthogonally diagonalizable. We prove this by supposing A is orthogonally diagonalizable, then:

$$\begin{aligned}
A &= PDP^{-1} \\
&= PDP^T \\
A^T &= (PDP^T)^T \\
&= \left(P^T\right)^T D^T P^T \\
&= PDP^T \\
&= A
\end{aligned}$$

We have a theorem that if A is a symmetric $n \times n$ matrix, then:

- A has n real eigenvalues counted using multiplicity
- $\dim \mathrm{Nul}(A - \lambda I) =$ multiplicity of λ in its characteristic polynomial
- Eigenspaces are mutually orthogonal

- A is orthogonally diagonalizable

Now, we may carry out a special decomposition. Say A is symmetric. We can write $A = PDP^{-1} = PDP^T$, where $P = \left[\ \vec{v}_1 \mid \dots \mid \vec{v}_n\ \right]$, where $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthogonal basis of eigenvectors. Then:

$$\left[\ \vec{v}_1 \mid \dots \mid \vec{v}_n\ \right] \begin{bmatrix} \lambda_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix} = \lambda_1 \vec{v}_1 \vec{v}_1^T + \dots + \lambda_n \vec{v}_n \vec{v}_n^T$$

Thus, $A = \lambda_1 \vec{v}_1 \vec{v}_1^T + \dots + \lambda_n \vec{v}_n \vec{v}_n^T$. This is called the spectral decomposition of A.

We claim that $\vec{u}_j \vec{u}_j^T$ is a projection matrix onto a 1-dimensional subspace (if $\|\vec{u}_j\| = 1$). The reason is that:

$$\begin{aligned} (\vec{u}_j \vec{u}_j^T)\vec{x} &= \vec{u}_j\left(\vec{u}_j^T \vec{x}\right) \\ &= \vec{u}_j(\vec{u}_j \cdot \vec{x}) \\ &= (\vec{u}_j \cdot \vec{x})\vec{u}_j \\ &= \mathrm{proj}_{\mathrm{span}(\vec{u}_j)} \vec{x} \end{aligned}$$

What this says is that A is a weighted combination of projections onto eigenvectors – this is a 'spectral decomposition' of a matrix.

25 October 24 Lecture

If A is a square matrix, we have eigenvalues – but what if A is not square? Supposed A is an $m \times n$ matrix, then $A^T A$ is an $n \times n$ matrix. We claim that $A^T A$ is symmetric (shown as $(A^T A)^T = A^T (A^T)^T = A^T A$). Thus $A^T A$ has n real eigenvalues. We claim that they are all nonnegative, and show this by saying λ is an eigenvalue – then $A^T A \vec{x} = \lambda \vec{x}$ for some $\vec{x} \in \mathbb{R}^n$. Then:

$$\begin{aligned} \|A\vec{x}\|^2 &= (A|vecx) \cdot (A\vec{x}) \\ &= (A\vec{x})^T A\vec{x} \\ &= (\vec{x}^T A^T)(A\vec{x}) \\ &= \vec{x}^T (A^T A\vec{x}) \\ &= \vec{x}^T \lambda \vec{x} \\ &= \lambda \|\vec{x}\|^2 \end{aligned}$$

Thus $\|A\vec{x}\|^2 = \lambda \|\vec{x}\|^2$, and $\lambda = \frac{\|A\vec{x}\|}{\|\vec{x}\|^2} \geq 0$.

We make the claim that the 0-eigenspace of $A^T A$ is $\mathrm{Nul}(A)$. We do this by supposing $\vec{x}$ is in 0-eigenspace, then $A^T A \vec{x} = 0$, and we know that $\|A\vec{x}\|^2 = \lambda \|\vec{x}\|^2 = 0$. Thus $A\vec{x} = 0$, and so $\vec{x} \in \mathrm{Nul}A$. $A^T A$ has $\dim(\mathrm{Nul}A)$ 0-eigenvalues. Thus, the number of positive eigenvalues of $A^T A$ is $\mathrm{rank}(A)$.

If we list the eigenvalues of A in decreasing order $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$ where $r = \mathrm{rank}(A)$. Then we suppose that we have $\{\vec{v}_1, \dots, \vec{v}_2\}$ that is an orthonormal set of eigenvectors for the positive eigenvalues of $A^T A$, then $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal basis of $\text{Col}(A)$. We show this by showing if $i \neq j$ then $A\vec{v}_i$ is orthogonal to $A\vec{v}_j$:

$$\begin{aligned}(A\vec{v}_i) \cdot (A\vec{v}_j) &= (A\vec{v}_i)^T A\vec{v}_j \\ &= \vec{v}_i^T A^T A\vec{v}_j \\ &= \vec{v}_i^T \lambda_j \vec{v}_j \\ &= \lambda_j \vec{v}_i \cdot \vec{v}_j \\ &= 0\end{aligned}$$

Now, since $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ are orthogonal and $r = \text{rank}(A)$, they are linearly independent, so they form an orthogonal basis of $\text{Col}(A)$.

We define the singular values of A as $\sigma_i = \sqrt{\lambda_i}$. Now, the purpose of this is if we can write $A = PDP^{-1}$ (when it is a square matrix), we would like to do something similar for when it is not a square matrix – what if it is $m \times n$?. So, now we make n $m \times n$ matrix σ:

$$\begin{bmatrix} \sigma_1 & \dots & 0 & \dots & 0 \\ 0 & \ddots & 0 & \dots & 0 \\ 0 & \dots & \sigma_r & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we can write $A = U\sigma V^T$, where U is an orthogonal $m \times m$ matrix and V is an orthogonal $n \times n$ matrix. This is called the singular value decomposition of A. We have $V = \left[\ \vec{v}_1 \mid \dots \mid \vec{v}_n\ \right]$. Then, if $\{\vec{v}_1, \dots, \vec{v}_r\}$ is an orthonormal collection of eigenvectors of $A^T A$ corresponding to $\lambda_1, \dots, \lambda_r$, then for $1 \leq i \leq r$, put $\vec{u}_i = \frac{A\vec{v}_i}{\|A\vec{v}_i\|}$, then $\{\vec{u}_1, \dots, \vec{u}_r\}$ is an orthonormal basis of $\text{Col}(A)$. Then we complete this to an orthonormal basis of $\{\vec{u}_1, \dots, \vec{u}_m\}$ of $\mathbb{R}^m$, then $U = \left[\ \vec{u}_1 \mid \dots \mid \vec{u}_m\ \right]$.

26 October 29 Lecture

We have a differential equation:

$$ay'' + by' + cy = 0$$

We guess:

$$\begin{aligned} y &= e^{rt} \\ y' &= re^{rt} \\ y'' &= r^2 e^{rt} \end{aligned}$$

Then:

$$e^{rt}\big(ar^2 + br + c\big) = 0$$

Now, there's two values of r that will work – and we say if y_1 and y_2 are solutions of the differential equation, then so is $c_1 y_1 + c_2 y_2$. Thus the solutions of the vector space is a vector space – in other words, there's a linear isomorphism from the solutions of the differential equation to $\mathbb{R}^2$. Thus if the auxiliary equation $ar^2 + br + c$ has two distinct real roots r_1 and r_2, then the solutions of the differential equation are any linear combination of $e^{r_1 t}, e^{r_2 t}$.

Now, we can uniquely determine a solution of the differential equation if we are **given** $y(0)$ and $y'(0)$ (i.e. initial conditions).

What happens if the auxiliary equation has a double zero r? One solution is still $y = e^{rt}$, and we claim the other is te^{rt}. We show this as follows. Take $ay'' + by' + c = 0$, assuming $a = 1$ (if not we can just divide both sides of the equation by a). Then, if we have a double root at r, then the equation must be $y'' - 2ry' + r^2 y = e^{\lambda t}\left(\lambda^2 - 2r\lambda + r^2\right) = e^{\lambda t}(\lambda - r)^2$. We now put $y = te^{rt}$, so that $y' = e^{rt} + rte^{rt}$, and $y'' = 2re^{rt} + r^2 te^{rt}$. Then:

$$\begin{aligned} y'' - 2ry' + r^2 y &= 2re^{rt} + r^2 te^{rt} - 2r\left(e^{rt} + rte^{rt}\right) + r^2 te^{rt} \\ &= 0 \end{aligned}$$

26.1 Wronskians

Say I is an interval in $\mathbb{R}$ ($a \le t \le b$). Then say y_1 and y_2 are two functions of t for $t \in I$, then we define the Wronskian as $W[y_1, y_2] = y_1 y_2' - y_2 y_1'$. We claim that if y_1 and y_2 are solutions of the differential equation $ay'' + by' + c = 0$, then $W[y_1, y_2] = Ce^{-\frac{b}{a}t}$ for some $C \in \mathbb{R}$.

We have $W = y_1 y_2' - y_2 y_1'$, and $W' = y_1' y_2' + y_1 y_2'' - y_2' y_1' - y_2 y_1'' = y_1 y_2'' - y_2 y_1''$:

$$\begin{aligned} aW' + bW &= a(y_1 y_2'' - y_2 y_1'') + b(y_1 y_2' - y_2 y_1') \\ &= y_1(ay_2'' + by_2' + cy_2) - y_2(ay_1'' + by_1' + cy_1) \\ &= 0 \end{aligned}$$

So $aW' + bW = 0$, then $W' + \frac{b}{a}W = 0$, then $W = Ce^{-\frac{b}{a}t}$ for some C (note $C \neq c$ from the differential equation). We define a pair of functions $y_1(t)$ and $y_2(t)$ as linearly dependent on I if one function is a constant times the other. Now, we claim that if $y_1(t)$ and $y_2(t)$ is linearly dependent, then $W[y_1, y_2] = 0$. Also, if $W[y_1, y_2] = 0$, then they are linearly dependent.

27 October 31 Lecture

We prove that if $W[y_1, y_2](\tau) = 0$ for some τ, y_1, y_2 satisfy $ay'' + by' + c = 0$, then y_1 and y_2 are linearly dependent. The idea is:

$$W[y_1, y_2](\tau) = Ce^{-\frac{b}{a}\tau} = 0$$

Thus $C = 0$, and $W[y_1, y_2] = 0$ for all t, and thus we can show y_1, y_2 are dependent.

Now, say $r = \alpha + i\beta$ is a root of $ar^2 + br + c = 0$. We know then that $r = \alpha - i\beta$ is also a solution. Thus we say the general solution would be $c_1 e^{(\alpha + i\beta)t} + c_2 e^{(\alpha - i\beta)t}$. Now, what does it mean to have $e^{i\beta t}$? By Taylor expansion, we can write $e^{\theta i} = \cos\theta + i\sin\theta$, so $e^{i\beta t} = \cos(\beta t) + i\sin(\beta t)$. Now, $e^{(a+i\beta)t} = e^{at}(\cos(\beta t) - i\sin(\beta t))$. We now note that if $y(t)$ is a complex-valued solution to $ay'' + by' + cy = 0$, then so are $\mathrm{Re}(y(t))$ and $\mathrm{Im}(y(t))$. Applying this to $y(t) = e^{(\alpha+i\beta)t} = e^{\alpha t}\cos(\beta t) + ie^{\alpha t}\sin(\beta t)$, we see that the general solution **real-valued** is thus:

$$y(t) = c_1 e^{\alpha t}\cos(\beta t) + c_2 e^{\alpha t}\sin(\beta t)$$

28 November 2 Lecture

We now introduce a differential equation with $ay'' + by' + cy = f(t)$. If y_1 and y_2 solve this, then $y_1 - y_2$ solves $ay'' + by' + cy = 0$. We have a theorem that if y_p is any particular solution of $ay'' + by' + cy = f(t)$, and y_h is the general solution of $ay'' + by' + cy = 0$, then $y_p + y_h$ solves $ay'' + by' + cy = f(t)$ as well. How do we find the particular solution? We make an intelligent guess. For example, $my'' + by' + Cy = F\cos(\beta t)$. We guess $y = c_1\cos(\beta t) + c_2\sin(\beta t)$:

$$\begin{aligned}
my'' + by' + Cy &= F\cos(\beta t) \\
= -mc_1\beta^2\cos(\beta t) - mc_2\beta^2\sin(\beta t) - c_1 b\beta\sin(\beta t) &+ c_2 b\beta\cos(\beta t) + c_1 C\cos(\beta t) + c_2 C\sin(\beta t) \\
= \cos(\beta t)\big(-mc_1\beta^2 + c_2 b\beta + c_1 C\big) &+ \sin(\beta t)\big(-mc_2\beta^2 - c_1 b\beta + c_2 C\big) \\
-mc_1\beta^2 + c_2 b\beta + c_1 C &= F \\
-mc_2\beta^2 - c_1 b\beta + c_2 C &= 0
\end{aligned}$$

This is solvable!

29 November 5 Lecture

We last had $ay'' + by' + cy = f(T)$. The general solution is $y = y_P + c_1 y_1 + c_2 y_2$ where y_p is any particular solution and y_1, y_2 are linearly independent solutions of $ay'' + by' + cy = 0$.

Say we have $y'' + y' + y = T^2$, so we guess $y = A_2T^2 + A_1T + A_0$, and $y' = 2A_2T + A_1$, and $y'' = 2A_2$. We plug this in to get:

$$\begin{aligned}
2A_2 + 2A_2T + A_1 + A_2T^2 + A_1T + A_0 &= T^2 \\
A_2 &= 1 \\
2A_2 + A_1 &= 0 \\
2A_2 + A_1 + A_0 &= 0
\end{aligned}$$

So we get:

$$\begin{aligned}
A_2 &= 1 \\
A_1 &= -2 \\
A_0 &= 0
\end{aligned}$$

Our particular solution is thus $y_p = T^2 - 2T$, and then we solve for $y'' + y' + y = 0$.

Now, say we have $y'' - 3y' + 2y = e^T$. We might guess $y = ce^T$, but one can quickly see this does not work out. Instead we guess:

$$\begin{aligned}
y &= cTe^T \\
y' &= c(e^T + Te^T) \\
y'' &= c(2e^T + Te^T) \\
e^T &= c(2e^T + Te^T) - 3c(e^T + Te^T) + 2cTe^T
\end{aligned}$$

$$\begin{aligned}
c - 3c + 2c &= 0 \\
2c - 3c &= 1 \\
-c &= 1
\end{aligned}$$

Thus the particular solution is $-Te^T$.

We have a theorem that if y_1 satisfies $ay''+by'+cy = f_1(T)$ and y_2 satisfies $ay''+by'+cy = f_2(T)$, then $y = k_1y_1 + k_2y_2$ satisfies $ay'' + by' + cy = k_1f_1(T) + k_2f_2(T)$ for $k_1, k_2 \in \mathbb{R}$. We prove this as follows:

$$\begin{aligned}
a(k_1y_1)'' + b(k_1y_1)' + ck_1y_1 &= k_1f_1 \\
a(k_2y_2)'' + b(k_2y_2)' + ck_2y_2 &= k_2f_2
\end{aligned}$$

Simply adding:

$$a(k_1y_1 + k_2y_2)'' + b(k_1y_1 + k_2y_2)' + c(k_1y_1 + k_2y_2) = k_1f_1 + k_2f_2$$

We have a theorem that given T_0, Y_0, Y_1, there's a unique solution to $ay'' + by' + c = f(T)$ with $y(T_0) = Y_0$ and $y'(T_0) = Y_1$, there is a unique solution.

To find a solution of $ay'' + by' + cy = P_m(T)e^{rT}$, where P_m is a polynomial of degree m, we guess $y_p = T^s(A_mT^m + \ldots + A_1T + A_0)e^{rT}$, where:

1. $s = 0$ if r is not a root of the auxiliary equation
2. $s = 1$ if r is a simple root of the auxiliary equation
3. $s = 2$ if r is a double root

Or, if we want to find a particular solution of $ay'' + by' + cy = P_m(T)e^{\alpha T}\cos(\beta T) + Q_n(T) + e^{\alpha T}\sin(\beta T)$, where $\beta \neq 0$, P_m is a polynomial of degree m and P_n is a polynomial of degree n. Then we guess:

$$y_p = T^s(A_KT^K + \ldots + A_1T + A_0)e^{\alpha T}\cos(\beta T) + T^s(\beta_kT_k + \ldots + \beta T + B_0)e^{\alpha T}\sin(\beta T)$$

1. $s = 0$ if $\alpha + i\beta$ is not a solution of the auxiliary equation
2. $s = 1$ if $\alpha + i\beta$ is a root of the auxiliary equation

30 November 7 Lecture

Now, instead of intelligent guessing we try to solve $ay'' + by' + cy = f(t)$ by varying parameters. Say y_1 and y_2 are linearly independent solutions of the homogenous equation $ay'' + by' + cy = 0$, then we guess a solution of $y_p = v_1(t)y_1(t) + v_2(t)y_2(t)$. We need to find v_1 and v_2 so that y_p solves the original differential equation. Now:

$$\begin{aligned} y &= v_1y_1 + v_2y_2 \\ y' &= v_1'y_1 + v_1y_1' + v_2'y + v_2y_2' \end{aligned}$$

y'' turns out to be very messy. To avoid this, we impose condition that $v_1'y_1 + v_2'y = 0$. Now:

$$\begin{aligned} y' &= v_1y_1' + v_2y_2' \\ y'' &= v_1'y_1' + v_1y_1'' + v_2'y_2' + v_2y_2'' \end{aligned}$$

Plugging this in to the original equation:

$$\begin{aligned} a(v_1'y_1' + v_1y_1'' + v_2'y_2' + v_2y_2'') + b(v_1y_1' + v_2y_2') + c(v_1y_1 + v_2y_2) &= f \\ v_1(ay_1'' + by_1' + cy_1) + v_2(ay_2'' + by_2' + cy_2) + av_1'y_1' + av_2'y_2' &= f \end{aligned}$$

Notice $ay_1'' + by_1' + cy_1 = 0$ and the same for y_2. We now have two equations:

$$\begin{aligned} v_1'y_1' + v_2'y_2' &= \frac{f}{a} \\ y_1v_1' + y_2v_2' &= 0 \end{aligned}$$

We can solve these two equations. So we solve for v_1' and v_2', integrate, and then the general solution is $v_1y_1 + v_2y_2 + c_1y_1 + c_2y_2$ for some constants $c_1, c_2 \in \mathbb{R}$.

Say we have $y'' + 2y' + 3y = 0$. We put $x_1(t) = y(t)$ and $x_2(t) = y'(t)$. Now:

$$\begin{cases} x_1' &= y' = x_2 \\ x_2' &= y'' = -2y' - 3y = -2x_2 - 3x_1 \end{cases}$$

Or:

$$\begin{cases} x_1' &= x_2 \\ x_2' &= -3x_1 - 2x_2 \end{cases}$$

Now, we can say $\vec{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$ and $\vec{x}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} x_2 \\ -3x_1 - 2x_1 \end{bmatrix}$. Now:

$$\vec{x}'(t) = \begin{bmatrix} 0 & 1 \\ -3 & -2 \end{bmatrix} \vec{x}(t)$$

Thus, to find a solution, it's enough to specify $t(t_0)$ and $y'(t_0)$ because this tells us what $\vec{x})(t_0)$ is.

Take a more complex example.

$$\begin{aligned} y''' + 2y'' + 3y' + 4y &= 0 \\ x_1 &= y \\ x_2 &= y' \\ x_3 &= y'' \\ x_1' &= x_2 \\ x_2' &= x_3 \\ x_3' &= -2x_3 - 3x_2 - 4x_1 \end{aligned}$$

Now, we put $\vec{x} = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$, then:

$$\vec{x}' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -3 & -2 \end{bmatrix} \vec{x}$$

So, we can say everything we've discussed so far is a special case of a first order linear system: $\vec{x}'(t) = A(t)\vec{x}(t) + \vec{f}(t)$, where $A(t)$ is a $n \times n$ matrix whose entries are functions of T. This is homogeneous if $\vec{f}(t) = 0$.

Now, take $x_1' = e^t x_1 + \sin(t)x_2 + 3t$, and $x_2' = x_1 + 4Tx_2 + 5t^3$. Now, we can write:

$$\vec{x}'(t) = \begin{bmatrix} e^t & \sin(t) \\ 1 & 4t \end{bmatrix} \vec{x} + \begin{bmatrix} 3t \\ 5t^3 \end{bmatrix}$$

Now, theoretically if we have $\vec{x} = A\vec{x} + f$, then we say $\vec{f} = \vec{x}' - A\vec{x}$. We define $L(\vec{x}) = \vec{x}' - A\vec{x}$ for $\vec{x}$ is an n-vector of functions. We note that L is linear (i.e. $L(c_1\vec{x}_1 + c_2\vec{x}_2) = c_1L(\vec{x}_1) + c_2L(\vec{x}_2)$. Now, i4 $\vec{x}_p$ is a particular solution of $\vec{x}' = A\vec{x} + f$ then the general solution is $\vec{x}_p + \vec{x}_n$, where $\vec{x}_n' = A\vec{x}_n$ ($L(\vec{x}_n) = 0$). The kernel of L is the set of all $\vec{x}$ so that $\vec{x}' = A\vec{x}$, so it is a vector space of dimension n. Then, if $\{\vec{x}_1, \ldots, \vec{x}_n\}$ is a fundamental solution set (i.e. they are a basis for the kernel of L), then the general solution for $\vec{x}' = A\vec{x} + f$ is:

$$\begin{aligned} \vec{x} &= \vec{x}_p + c_1\vec{x}_1 + \ldots + c_n\vec{x}_n \\ &= \left[\begin{array}{c|c|c} \vec{x}_1(t) & \ldots & \vec{x}_n(t) \end{array}\right] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \vec{x}_p \end{aligned}$$

Now, we define the Wronskian $W[\vec{x}_1, \ldots, \vec{x}_n] = \det \left[\begin{array}{c|c|c} \vec{x}_1 & \ldots & \vec{x}_n \end{array}\right]$. Now, we note if $\vec{x}_1, \ldots, \vec{x}_n$ is linearly dependent then the Wronskian is 0.

31 November 14 Lecture

We now study the homogeneous equation $\vec{x}' = A\vec{x}$ where A is an $n \times n$ matrix with constant entries. We guess $\vec{x} = e^{rt}\vec{u}$ and get:

$$\begin{aligned} re^{rt}\vec{u} &= Ae^{rt}\vec{u} \\ r\vec{u} &= A\vec{u} \end{aligned}$$

Thus we see $\vec{x} = e^{rt}\vec{u}$ solves the equation if r is an eigenvalue of A and $\vec{u}$ is its corresponding eigenvector. Thus, if A has n linearly independent eigenvalues, then $\{e^{r_1t}\vec{u}_1, \ldots, e^{r_nt}\vec{u}_n\}$ forms a fundamental solution set.

32 November 26 Lecture

33 November 28 Lecture

If we define a vector space V to be the set of all functions f such that $f(x+2T) = f(x)$ for all x (where T is the period), we define a linear operator $L(f) = f''$ and the inner product to be $\langle f, g\rangle = \int_{-T}^{T} f(x)g(x)\,\mathrm{d}x$. Then, L is symmetric (i.e. $\langle L(f), g\rangle = \langle f, L(g)\rangle$). Furthermore, the eigenvalues of L are $-\left(\frac{n\pi}{T}\right)^2$, $n = 0, 1, 2, \ldots$. The corresponding eigenvectors are then $\sin\left(\frac{n\pi x}{T}\right), \cos\left(\frac{n\pi x}{T}\right)$. We know eigenvectors corresponding to distinct eigenvalues are orthogonal (i.e. their inner product is 0).

Now, we imagine an arbitrary function f – we can decompose it into the various eigenvectors of L so that:

$$f = \sum_i \frac{\langle f, u_i\rangle}{\langle u_i, u_i\rangle} u_i$$

Here, $\{u_i\}$ is our orthonormal basis. However, we note that $\langle \sin\left(\frac{n\pi x}{T}\right)), \sin\left(\frac{n\pi x}{T}\right)\rangle = T$ and $\langle \cos\left(\frac{n\pi x}{T}\right), \cos\left(\frac{n\pi x}{T}\right)\rangle = T$ (we show this by noting that these two integrals are equal to one another, and their sum is $2T$). Thus,

$$f = \frac{\langle f, 1\rangle}{2T} 1 + \frac{1}{T}\sum_{n=1}^{\infty}\left(\langle f, \sin\left(\frac{n\pi x}{T}\right)\rangle \sin\left(\frac{n\pi x}{T}\right) + \langle f, \cos\left(\frac{n\pi x}{T}\right)\rangle \cos\left(\frac{n\pi x}{T}\right)\right)$$

Take for example, $f = |x|$ on $-\pi \le x \le \pi$ (we imagine this to be periodic, so it is a sawtooth pattern). Now:

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} |x| \sin(nx)\,\mathrm{d}x$$

Since the integrand is odd:

$$\begin{aligned}
b_n &= 0 \\
a_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |x|\,\mathrm{d}x \\
&= \pi \\
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} |x| \cos(nx)\,\mathrm{d}x \\
&= -\frac{2}{\pi n^2}(1 - (-1)^n) \\
|x| &+ \frac{\pi}{2} - \sum_{n=1}^{\infty} \frac{2}{\pi n^2}(1 - (-1)^n)
\end{aligned}$$