The geometry of linear equations

The fundamental problem of linear algebra is to solve n linear equations in n unknowns; for example:

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3. \end{aligned}$$

In this first lecture on linear algebra we view this problem in three ways.

The system above is two dimensional ($n = 2$). By adding a third variable z we could expand it to three dimensions.

Row Picture

Plot the points that satisfy each equation. The intersection of the plots (if they do intersect) represents the solution to the system of equations. Looking at Figure 1 we see that the solution to this system of equations is $x = 1, y = 2$.

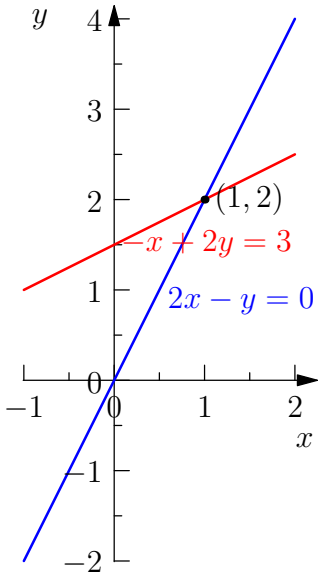

Figure 1: The lines $2x - y = 0$ and $-x + 2y = 3$ intersect at the point $(1, 2)$.

We plug this solution in to the original system of equations to check our work:

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

The solution to a three dimensional system of equations is the common point of intersection of three planes (if there is one).

Column Picture

In the column picture we rewrite the system of linear equations as a single equation by turning the coefficients in the columns of the system into vectors:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

Given two vectors $\mathbf{c}$ and $\mathbf{d}$ and scalars x and y, the sum $x\mathbf{c} + y\mathbf{d}$ is called a *linear combination* of $\mathbf{c}$ and $\mathbf{d}$. Linear combinations are important throughout this course.

Geometrically, we want to find numbers x and y so that x copies of vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ added to y copies of vector $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ equals the vector $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$. As we see from Figure 2, $x = 1$ and $y = 2$, agreeing with the row picture in Figure 2.

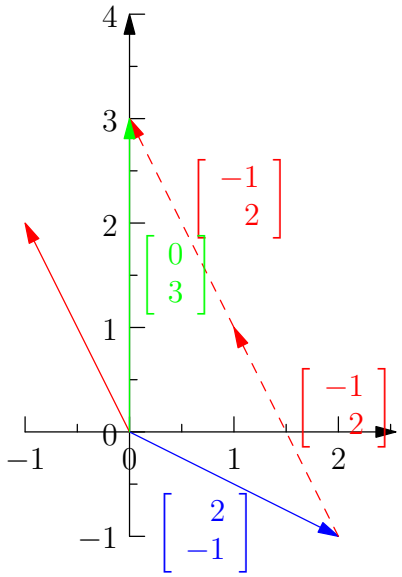

Figure 2: A linear combination of the column vectors equals the vector $\mathbf{b}$.

In three dimensions, the column picture requires us to find a linear combination of three 3-dimensional vectors that equals the vector $\mathbf{b}$.

Matrix Picture

We write the system of equations

$$\begin{aligned} 2x - y &= 0 \\ -x + 2y &= 3 \end{aligned}$$

as a single equation by using matrices and vectors:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}.$$

The matrix $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ is called the *coefficient matrix*. The vector $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is the vector of unknowns. The values on the right hand side of the equations form the vector $\mathbf{b}$:

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

The three dimensional matrix picture is very like the two dimensional one, except that the vectors and matrices increase in size.

Matrix Multiplication

How do we multiply a matrix A by a vector $\mathbf{x}$?

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \quad ?$$

One method is to think of the entries of $\mathbf{x}$ as the coefficients of a linear combination of the column vectors of the matrix:

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}.$$

This technique shows that $A\mathbf{x}$ is a linear combination of the columns of A.

You may also calculate the product $A\mathbf{x}$ by taking the dot product of each row of A with the vector $\mathbf{x}$:

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 5 \cdot 2 \\ 1 \cdot 1 + 3 \cdot 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}.$$

Linear Independence

In the column and matrix pictures, the right hand side of the equation is a vector $\mathbf{b}$. Given a matrix A, can we solve:

$$A\mathbf{x} = \mathbf{b}$$

for every possible vector $\mathbf{b}$? In other words, do the linear combinations of the column vectors fill the xy-plane (or space, in the three dimensional case)?

If the answer is "no", we say that A is a *singular matrix*. In this singular case its column vectors are *linearly dependent*; all linear combinations of those vectors lie on a point or line (in two dimensions) or on a point, line or plane (in three dimensions). The combinations don't fill the whole space.

An overview of key ideas

This is an overview of linear algebra given at the start of a course on the mathematics of engineering.

Linear algebra progresses from vectors to matrices to subspaces.

Vectors

What do you do with vectors? Take combinations.

We can multiply vectors by scalars, add, and subtract. Given vectors **u**, **v** and **w** we can form the *linear combination* $x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = \mathbf{b}$.

An example in $\mathbb{R}^3$ would be:

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The collection of all multiples of **u** forms a line through the origin. The collection of all multiples of **v** forms another line. The collection of all combinations of **u** and **v** forms a plane. Taking *all combinations* of some vectors creates a *subspace*.

We could continue like this, or we can use a matrix to add in all multiples of **w**.

Matrices

Create a matrix A with vectors **u**, **v** and **w** in its columns:

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

The product:

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}$$

equals the sum $x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = \mathbf{b}$. The product of a matrix and a vector is a combination of the columns of the matrix. (This particular matrix A is a *difference matrix* because the components of $A\mathbf{x}$ are differences of the components of that vector.)

When we say $x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = \mathbf{b}$ we're thinking about multiplying numbers by vectors; when we say $A\mathbf{x} = \mathbf{b}$ we're thinking about multiplying a matrix (whose columns are **u**, **v** and **w**) by the numbers. The calculations are the same, but our perspective has changed.

For any input vector $\mathbf{x}$, the output of the operation "multiplication by A" is some vector $\mathbf{b}$:

$$A \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}.$$

A deeper question is to start with a vector $\mathbf{b}$ and ask "for what vectors $\mathbf{x}$ does $A\mathbf{x} = \mathbf{b}$?" In our example, this means solving three equations in three unknowns. Solving:

$$A\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

is equivalent to solving:

$$\begin{aligned} x_1 &= b_1 \\ x_2 - x_1 &= b_2 \\ x_3 - x_2 &= b_3. \end{aligned}$$

We see that $x_1 = b_1$ and so x_2 must equal $b_1 + b_2$. In vector form, the solution is:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_1 + b_2 \\ b_1 + b_2 + b_3 \end{bmatrix}.$$

But this just says:

$$\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

or $\mathbf{x} = A^{-1}\mathbf{b}$. If the matrix A is invertible, we can multiply on both sides by A^{-1} to find the unique solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$. We might say that A represents a transform $\mathbf{x} \to \mathbf{b}$ that has an inverse transform $\mathbf{b} \to \mathbf{x}$.

In particular, if $\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ then $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

The second example has the same columns $\mathbf{u}$ and $\mathbf{v}$ and replaces column vector $\mathbf{w}$:

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

Then:

$$C\mathbf{x} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

and our system of three equations in three unknowns becomes circular.

Where before $A\mathbf{x} = \mathbf{0}$ implied $\mathbf{x} = \mathbf{0}$, there are non-zero vectors $\mathbf{x}$ for which $C\mathbf{x} = \mathbf{0}$. For any vector $\mathbf{x}$ with $x_1 = x_2 = x_3$, $C\mathbf{x} = \mathbf{0}$. This is a significant difference; we can't multiply both sides of $C\mathbf{x} = \mathbf{0}$ by an inverse to find a non-zero solution $\mathbf{x}$.

The system of equations encoded in $C\mathbf{x} = \mathbf{b}$ is:

$$\begin{aligned} x_1 - x_3 &= b_1 \\ x_2 - x_1 &= b_2 \\ x_3 - x_2 &= b_3. \end{aligned}$$

If we add these three equations together, we get:

$$0 = b_1 +_2 + b_3.$$

This tells us that $C\mathbf{x} = \mathbf{b}$ has a solution $\mathbf{x}$ only when the components of $\mathbf{b}$ sum to 0. In a physical system, this might tell us that the system is stable as long as the forces on it are balanced.

Subspaces

Geometrically, the columns of C lie in the same plane (they are *dependent*; the columns of A are *independent*). There are many vectors in $\mathbb{R}^3$ which do not lie in that plane. Those vectors cannot be written as a linear combination of the columns of C and so correspond to values of $\mathbf{b}$ for which $C\mathbf{x} = \mathbf{b}$ has no solution $\mathbf{x}$. The linear combinations of the columns of C form a two dimensional *subspace* of $\mathbb{R}^3$.

This plane of combinations of $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ can be described as "all vectors $C\mathbf{x}$". But we know that the vectors $\mathbf{b}$ for which $C\mathbf{x} = \mathbf{b}$ satisfy the condition $b_1 + b_2 + b_3 = 0$. So the plane of all combinations of $\mathbf{u}$ and $\mathbf{v}$ consists of all vectors whose components sum to 0.

If we take all combinations of:

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \text{ and } \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

we get the entire space $\mathbb{R}^3$; the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^3$. We say that $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ form a *basis* for $\mathbb{R}^3$.

A *basis* for $\mathbb{R}^n$ is a collection of n independent vectors in $\mathbb{R}^n$. Equivalently, a basis is a collection of n vectors whose combinations cover the whole space. Or, a collection of vectors forms a basis whenever a matrix which has those vectors as its columns is invertible.

A *vector space* is a collection of vectors that is closed under linear combinations. A *subspace* is a vector space inside another vector space; a plane through the origin in $\mathbb{R}^3$ is an example of a subspace. A subspace could be equal to the space it's contained in; the smallest subspace contains only the zero vector.

The subspaces of $\mathbb{R}^3$ are:

- the origin,
- a line through the origin,
- a plane through the origin,
- all of $\mathbb{R}^3$.

Conclusion

When you look at a matrix, try to see "what is it doing?"

Matrices can be rectangular; we can have seven equations in three unknowns. Rectangular matrices are not invertible, but the symmetric, square matrix $A^T A$ that often appears when studying rectangular matrices may be invertible.

Elimination with matrices

Method of Elimination

Elimination is the technique most commonly used by computer software to solve systems of linear equations. It finds a solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$ whenever the matrix A is invertible. In the example used in class,

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}.$$

The number 1 in the upper left corner of A is called the *first pivot*. We recopy the first row, then multiply the numbers in it by an appropriate value (in this case 3) and subtract those values from the numbers in the second row. The first number in the second row becomes 0. We have thus *eliminated* the 3 in row 2 column 1.

The next step is to perform another elimination to get a 0 in row 3 column 1; here this is already the case.

The *second pivot* is the value 2 which now appears in row 2 column 2. We find a multiplier (in this case 2) by which we multiply the second row to eliminate the 4 in row 3 column 2. The *third pivot* is then the 5 now in row 3 column 3.

We started with an invertible matrix A and ended with an *upper triangular* matrix U; the lower left portion of U is filled with zeros. Pivots 1, 2, 5 are on the diagonal of U.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} \longrightarrow U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

We repeat the multiplications and subtractions with the vector $\mathbf{b} = \begin{bmatrix} 2 \\ 12 \\ 2 \end{bmatrix}$. For example, we multiply the 2 in the first position by 3 and subtract from 12 to get 6 in the second position. When calculating by hand we can do this efficiently by *augmenting* the matrix A, appending the vector $\mathbf{b}$ as a fourth or final column. The method of elimination transforms the equation $A\mathbf{x} = \mathbf{b}$ into a new equation $U\mathbf{x} = \mathbf{c}$. In the example above, $U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$ comes from A and $\mathbf{c} = \begin{bmatrix} 2 \\ 6 \\ -10 \end{bmatrix}$ comes from $\mathbf{b}$.

The equation $U\mathbf{x} = c$ is easy to solve by *back substitution*; in our example, $z = -2$, $y = 1$ and $x = 2$. This is also a solution to the original system $A\mathbf{x} = \mathbf{b}$.

The *determinant* of U is the product of the pivots. We will see this again.

Pivots may not be 0. If there is a zero in the pivot position, we must exchange that row with one below to get a non-zero value in the pivot position.

If there is a zero in the pivot position and no non-zero value below it, then the matrix A is not invertible. Elimination can not be used to find a unique solution to the system of equations – it doesn't exist.

Elimination Matrices

The product of a matrix (3x3) and a column vector (3x1) is a column vector (3x1) that is a linear combination of the columns of the matrix.

The product of a row (1x3) and a matrix (3x3) is a row (1x3) that is a linear combination of the rows of the matrix.

We can subtract 3 times row 1 of matrix A from row 2 of A by calculating the matrix product:

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}.$$

The *elimination matrix* used to eliminate the entry in row m column n is denoted E_{mn}. The calculation above took us from A to $E_{21}A$. The three elimination steps leading to U were: $E_{32}(E_{31}(E_{21}A)) = U$, where $E_{31} = I$. Thus $E_{32}(E_{21}A) = U$.

Matrix multiplication is *associative*, so we can also write $(E_{32}E_{21})A = U$. The product $E_{32}E_{21}$ tells us how to get from A to U. The *inverse* of the matrix $E_{32}E_{21}$ tells us how to get from U to A.

If we solve $U\mathbf{x} = EA\mathbf{x} = E\mathbf{b}$, then it is also true that $A\mathbf{x} = \mathbf{b}$. This is why the method of elimination works: all steps can be reversed.

A *permutation matrix* exchanges two rows of a matrix; for example,

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

The first and second rows of the matrix PA are the second and first rows of the matrix A. The matrix P is constructed by exchanging rows of the identity matrix.

To exchange the columns of a matrix, multiply on the right (as in AP) by a permutation matrix.

Note that matrix multiplication is not *commutative*: $PA \neq AP$.

Inverses

We have a matrix:

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which subtracts 3 times row 1 from row 2. To "undo" this operation we must add 3 times row 1 to row 2 using the inverse matrix:

$$E_{21}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In fact, $E_{21}^{-1} E_{21} = I$.

Lecture 3: Multiplication and inverse matrices

Matrix Multiplication

We discuss four different ways of thinking about the product $AB = C$ of two matrices. If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then C is an $m \times p$ matrix. We use c_{ij} to denote the entry in row i and column j of matrix C.

Standard (row times column)

The standard way of describing a matrix product is to say that c_{ij} equals the dot product of row i of matrix A and column j of matrix B. In other words, $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

Columns

The product of matrix A and column j of matrix B equals column j of matrix C. This tells us that the columns of C are combinations of columns of A.

Rows

The product of row i of matrix A and matrix B equals row i of matrix C. So the rows of C are combinations of rows of B.

Column times row

A column of A is an $m \times 1$ vector and a row of B is a $1 \times p$ vector. Their product is a matrix:

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}.$$

The columns of this matrix are multiples of the column of A and the rows are multiples of the row of B. If we think of the entries in these rows as the coordinates $(2, 12)$ or $(3, 18)$ or $(4, 24)$, all these points lie on the same line; similarly for the two column vectors. Later we'll see that this is equivalent to saying that the *row space* of this matrix is a single line, as is the *column space*.

The product of A and B is the sum of these "column times row" matrices:

$$AB = \sum_{k=1}^{n} \begin{bmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{bmatrix} \begin{bmatrix} b_{k1} & \cdots & b_{kn} \end{bmatrix}.$$

Blocks

If we subdivide A and B into blocks that match properly, we can write the product $AB = C$ in terms of products of the blocks:

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}.$$

Here $C_1 = A_1B_1 + A_2B_3$.

Inverses

Square matrices

If A is a square matrix, the most important question you can ask about it is whether it has an inverse A^{-1}. If it does, then $A^{-1}A = I = AA^{-1}$ and we say that A is *invertible* or *nonsingular*.

If A is *singular* – i.e. A does not have an inverse – its determinant is zero and we can find some non-zero vector $\mathbf{x}$ for which $A\mathbf{x} = 0$. For example:

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In this example, three times the first column minus one times the second column equals the zero vector; the two column vectors lie on the same line.

Finding the inverse of a matrix is closely related to solving systems of linear equations:

$$\underset{A}{\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}} \quad \underset{A^{-1}}{\begin{bmatrix} a & c \\ b & d \end{bmatrix}} \quad = \quad \underset{I}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

can be read as saying "A times column j of A^{-1} equals column j of the identity matrix". This is just a special form of the equation $A\mathbf{x} = \mathbf{b}$.

Gauss-Jordan Elimination

We can use the method of elimination to solve two or more linear equations at the same time. Just augment the matrix with the whole identity matrix I:

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array}\right]$$

(Once we have used Gauss' elimination method to convert the original matrix to upper triangular form, we go on to use Jordan's idea of eliminating entries in the upper right portion of the matrix.)

$$A^{-1} = \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix}.$$

As in the last lecture, we can write the results of the elimination method as the product of a number of elimination matrices E_{ij} with the matrix A. Letting E be the product of all the E_{ij}, we write the result of this Gauss-Jordan elimination using block matrices: $E[\,A \mid I\,] = [\,I \mid E\,]$. But if $EA = I$, then $E = A^{-1}$.

Factorization into $A = LU$

One goal of today's lecture is to understand Gaussian elimination in terms of matrices; to find a matrix L such that $A = LU$. We start with some useful facts about matrix multiplication.

Inverse of a product

The inverse of a matrix product AB is $B^{-1}A^{-1}$.

Transpose of a product

We obtain the *transpose* of a matrix by exchanging its rows and columns. In other words, the entry in row i column j of A is the entry in row j column i of A^T.

The transpose of a matrix product AB is $B^T A^T$. For any invertible matrix A, the inverse of A^T is $\left(A^{-1}\right)^T$.

$A = LU$

We've seen how to use elimination to convert a suitable matrix A into an upper triangular matrix U. This leads to the factorization $A = LU$, which is very helpful in understanding the matrix A.

Recall that (when there are no row exchanges) we can describe the elimination of the entries of matrix A in terms of multiplication by a succession of elimination matrices E_{ij}, so that $A \to E_{21}A \to E_{31}E_{21}A \to \cdots \to U$. In the two by two case this looks like:

$$\underset{E_{21}}{\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}} \quad \underset{A}{\begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}} = \underset{U}{\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}}.$$

We can convert this to a factorization $A = LU$ by "canceling" the matrix E_{21}; multiply by its inverse to get $E_{21}^{-1}E_{21}A = E_{21}^{-1}U$.

$$\underset{A}{\begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}} \quad \underset{U}{\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}}.$$

The matrix U is upper triangular with pivots on the diagonal. The matrix L is *lower triangular* and has ones on the diagonal. Sometimes we will also want to factor out a diagonal matrix whose entries are the pivots:

$$\underset{A}{\begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}} \quad \underset{D}{\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}} \quad \underset{U'}{\begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}}.$$

In the three dimensional case, if $E_{32}E_{31}E_{21}A = U$ then $A = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1}U = LU$.

For example, suppose E_{31} is the identity matrix and E_{32} and E_{21} are as shown below:

$$\underset{E_{32}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix}} \quad \underset{E_{21}}{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad = \quad \underset{E}{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix}}.$$

The 10 in the lower left corner arises because we subtracted twice the first row from the second row, then subtracted five times the new second row from the third.

The factorization $A = LU$ is preferable to the statement $EA = U$ because the combination of row subtractions does not have the effect on L that it did on E. Here $L = E^{-1} = E_{21}^{-1}E_{32}^{-1}$:

$$\underset{E_{21}^{-1}}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{E_{32}^{-1}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}} \quad = \quad \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}}.$$

Notice the 0 in row three column one of $L = E^{-1}$, where E had a 10. If there are no row exchanges, the multipliers from the elimination matrices are copied directly into L.

How expensive is elimination?

Some applications require inverting very large matrices. This is done using a computer, of course. How hard will the computer have to work? How long will it take?

When using elimination to find the factorization $A = LU$ we just saw that we can build L as we go by keeping track of row subtractions. We have to remember L and (the matrix which will become) U; we don't have to store A or E_{ij} in the computer's memory.

How many operations does the computer perform during the elimination process for an $n \times n$ matrix? A typical operation is to multiply one row and then subtract it from another, which requires on the order of n operations. There are n rows, so the total number of operations used in eliminating entries in the first column is about n^2. The second row and column are shorter; that product costs about $(n-1)^2$ operations, and so on. The total number of operations needed to factor A into LU is on the order of n^3:

$$1^2 + 2^2 + \cdots + (n-1)^2 + n^2 = \sum_{i=1}^{n} i^2 \approx \int_0^n x^2\,dx = \frac{1}{3}n^3.$$

While we're factoring A we're also operating on $\mathbf{b}$. That costs about n^2 operations, which is hardly worth counting compared to $\frac{1}{3}n^3$.

Row exchanges

What if there are row exchanges? In other words, what happens if there's a zero in a pivot position?

To swap two rows, we multiply on the left by a permutation matrix. For example,

$$P_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

swaps the first and second rows of a 3×3 matrix. The inverse of any permutation matrix P is $P^{-1} = P^T$.

There are $n!$ different ways to permute the rows of an $n \times n$ matrix (including the permutation that leaves all rows fixed) so there are $n!$ permutation matrices. These matrices form a *multiplicative group*.

Transposes, permutations, spaces R^n

In this lecture we introduce vector spaces and their subspaces.

Permutations

Multiplication by a permutation matrix P swaps the rows of a matrix; when applying the method of elimination we use permutation matrices to move zeros out of pivot positions. Our factorization $A = LU$ then becomes $PA = LU$, where P is a permutation matrix which reorders any number of rows of A. Recall that $P^{-1} = P^T$, i.e. that $P^T P = I$.

Transposes

When we take the transpose of a matrix, its rows become columns and its columns become rows. If we denote the entry in row i column j of matrix A by A_{ij}, then we can describe A^T by: $\left(A^T\right)_{ij} = A_{ji}$. For example:

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

A matrix A is *symmetric* if $A^T = A$. Given any matrix R (not necessarily square) the product $R^T R$ is always symmetric, because $\left(R^T R\right)^T = R^T \left(R^T\right)^T = R^T R$. (Note that $\left(R^T\right)^T = R$.)

Vector spaces

We can add vectors and multiply them by numbers, which means we can discuss *linear combinations* of vectors. These combinations follow the rules of a *vector space*.

One such vector space is $\mathbb{R}^2$, the set of all vectors with exactly two real number components. We depict the vector $\begin{bmatrix} a \\ b \end{bmatrix}$ by drawing an arrow from the origin to the point (a, b) which is a units to the right of the origin and b units above it, and we call $\mathbb{R}^2$ the "$x - y$ plane".

Another example of a space is $\mathbb{R}^n$, the set of (column) vectors with n real number components.

Closure

The collection of vectors with exactly two *positive* real valued components is *not* a vector space. The sum of any two vectors in that collection is again in the collection, but multiplying any vector by, say, -5, gives a vector that's not

in the collection. We say that this collection of positive vectors is *closed* under addition but not under multiplication.

If a collection of vectors is closed under linear combinations (i.e. under addition and multiplication by any real numbers), and if multiplication and addition behave in a reasonable way, then we call that collection a *vector space*.

Subspaces

A vector space that is contained inside of another vector space is called a *subspace* of that space. For example, take any non-zero vector $\mathbf{v}$ in $\mathbb{R}^2$. Then the set of all vectors $c\mathbf{v}$, where c is a real number, forms a subspace of $\mathbb{R}^2$. This collection of vectors describes a line through $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ in $\mathbb{R}^2$ and is closed under addition.

A line in $\mathbb{R}^2$ that does not pass through the origin is *not* a subspace of $\mathbb{R}^2$. Multiplying any vector on that line by 0 gives the zero vector, which does not lie on the line. Every subspace must contain the zero vector because vector spaces are closed under multiplication.

The subspaces of $\mathbb{R}^2$ are:

1. all of $\mathbb{R}^2$,
2. any line through $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and
3. the zero vector alone (Z).

The subspaces of $\mathbb{R}^3$ are:

1. all of $\mathbb{R}^3$,
2. any plane through the origin,
3. any line through the origin, and
4. the zero vector alone (Z).

Column space

Given a matrix A with columns in $\mathbb{R}^3$, these columns and all their linear combinations form a subspace of $\mathbb{R}^3$. This is the *column space* $C(A)$. If $A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}$, the column space of A is the plane through the origin in $\mathbb{R}^3$ containing $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$.

Our next task will be to understand the equation $A\mathbf{x} = b$ in terms of subspaces and the column space of A.

Column space and nullspace

In this lecture we continue to study subspaces, particularly the column space and nullspace of a matrix.

Review of subspaces

A vector space is a collection of vectors which is closed under linear combinations. In other words, for any two vectors $\mathbf{v}$ and $\mathbf{w}$ in the space and any two real numbers c and d, the vector $c\mathbf{v} + d\mathbf{w}$ is also in the vector space. A subspace is a vector space contained inside a vector space.

A plane P containing $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and a line L containing $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ are both subspaces of $\mathbb{R}^3$. The union $P \cup L$ of those two subspaces is generally not a subspace, because the sum of a vector in P and a vector in L is probably not contained in $P \cup L$. The intersection $S \cap T$ of two subspaces S and T is a subspace. To prove this, use the fact that both S and T are closed under linear combinations to show that their intersection is closed under linear combinations.

Column space of A

The *column space* of a matrix A is the vector space made up of all linear combinations of the columns of A.

Solving $A\mathbf{x} = \mathbf{b}$

Given a matrix A, for what vectors $\mathbf{b}$ does $A\mathbf{x} = \mathbf{b}$ have a solution $\mathbf{x}$?

$$\text{Let } A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}.$$

Then $A\mathbf{x} = \mathbf{b}$ does not have a solution for every choice of $\mathbf{b}$ because solving $A\mathbf{x} = \mathbf{b}$ is equivalent to solving four linear equations in three unknowns. If there is a solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$, then $\mathbf{b}$ must be a linear combination of the columns of A. Only three columns cannot fill the entire four dimensional vector space – some vectors $\mathbf{b}$ cannot be expressed as linear combinations of columns of A.

Big question: what $\mathbf{b}$'s allow $A\mathbf{x} = \mathbf{b}$ to be solved?

A useful approach is to choose $\mathbf{x}$ and find the vector $\mathbf{b} = A\mathbf{x}$ corresponding to that solution. The components of $\mathbf{x}$ are just the coefficients in a linear combination of columns of A.

The system of linear equations $A\mathbf{x} = \mathbf{b}$ is *solvable* exactly when $\mathbf{b}$ is a vector in the *column space* of A.

For our example matrix A, what can we say about the column space of A? Are the columns of A *independent*? In other words, does each column contribute something new to the subspace?

The third column of A is the sum of the first two columns, so does not add anything to the subspace. The column space of our matrix A is a two dimensional subspace of $\mathbb{R}^4$.

Nullspace of A

The *nullspace* of a matrix A is the collection of all solutions $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ to the equation $A\mathbf{x} = 0$.

The column space of the matrix in our example was a subspace of $\mathbb{R}^4$. The nullspace of A is a subspace of $\mathbb{R}^3$. To see that it's a vector space, check that any sum or multiple of solutions to $A\mathbf{x} = \mathbf{0}$ is also a solution: $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0}$ and $A(c\mathbf{x}) = cA\mathbf{x} = c(\mathbf{0})$.

In the example:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

the nullspace $N(A)$ consists of all multiples of $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$; column 1 plus column 2 minus column 3 equals the zero vector. This nullspace is a line in $\mathbb{R}^3$.

Other values of b

The solutions to the equation:

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

do not form a subspace. The zero vector is not a solution to this equation. The set of solutions forms a line in $\mathbb{R}^3$ that passes through the points $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$ but not $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Solving $A\mathbf{x} = \mathbf{0}$: pivot variables, special solutions

We have a definition for the column space and the nullspace of a matrix, but how do we compute these subspaces?

Computing the nullspace

The *nullspace* of a matrix A is made up of the vectors $\mathbf{x}$ for which $A\mathbf{x} = \mathbf{0}$.

Suppose:

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}.$$

(Note that the columns of this matrix A are not independent.) Our algorithm for computing the nullspace of this matrix uses the method of elimination, despite the fact that A is not invertible. We don't need to use an augmented matrix because the right side (the vector $\mathbf{b}$) is $\mathbf{0}$ in this computation.

The row operations used in the method of elimination don't change the solution to $A\mathbf{x} = \mathbf{b}$ so they don't change the nullspace. (They do affect the column space.)

The first step of elimination gives us:

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \quad \longrightarrow \quad \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix}.$$

We don't find a pivot in the second column, so our next pivot is the 2 in the third column of the second row:

$$\begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 2 & 4 \end{bmatrix} \quad \longrightarrow \quad \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

The matrix U is in *echelon* (staircase) form. The third row is zero because row 3 was a linear combination of rows 1 and 2; it was eliminated.

The *rank* of a matrix A equals the number of pivots it has. In this example, the rank of A (and of U) is 2.

Special solutions

Once we've found U we can use back-substitution to find the solutions $\mathbf{x}$ to the equation $U\mathbf{x} = \mathbf{0}$. In our example, columns 1 and 3 are *pivot columns* containing pivots, and columns 2 and 4 are *free columns*. We can assign any value to x_2 and x_4; we call these *free variables*. Suppose $x_2 = 1$ and $x_4 = 0$. Then:

$$2x_3 + 4x_4 = 0 \quad \Longrightarrow \quad x_3 = 0$$

and:

$$x_1 + 2x_2 + 2x_3 + 2x_4 = 0 \quad \Longrightarrow \quad x_1 = -2.$$

So one solution is $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ (because the second column is just twice the first column). Any multiple of this vector is in the nullspace.

Letting a different free variable equal 1 and setting the other free variables equal to zero gives us other vectors in the nullspace. For example:

$$\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$$

has $x_4 = 1$ and $x_2 = 0$. The nullspace of A is the collection of all linear combinations of these "special solution" vectors.

The rank r of A equals the number of pivot columns, so the number of free columns is $n - r$: the number of columns (variables) minus the number of pivot columns. This equals the number of special solution vectors and the dimension of the nullspace.

Reduced row echelon form

By continuing to use the method of elimination we can convert U to a matrix R in *reduced row echelon form* (rref form), with pivots equal to 1 and zeros above and below the pivots.

$$U = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 & -2 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R.$$

By exchanging some columns, R can be rewritten with a copy of the identity matrix in the upper left corner, possibly followed by some free columns on the right. If some rows of A are linearly dependent, the lower rows of the matrix R will be filled with zeros:

$$R = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}.$$

(Here I is an r by r square matrix.)

If N is the *nullspace matrix* $N = \begin{bmatrix} -F \\ I \end{bmatrix}$ then $RN = 0$. (Here I is an $n - r$ by $n - r$ square matrix and 0 is an m by $n - r$ matrix.) The columns of N are the special solutions.

Solving $A\mathbf{x} = \mathbf{b}$: row reduced form R

When does $A\mathbf{x} = \mathbf{b}$ have solutions $\mathbf{x}$, and how can we describe those solutions?

Solvability conditions on b

We again use the example:

$$A = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix}.$$

The third row of A is the sum of its first and second rows, so we know that if $A\mathbf{x} = \mathbf{b}$ the third component of $\mathbf{b}$ equals the sum of its first and second components. If $\mathbf{b}$ does not satisfy $b_3 = b_1 + b_2$ the system has no solution. If a combination of the rows of A gives the zero row, then the same combination of the entries of $\mathbf{b}$ must equal zero.

One way to find out whether $A\mathbf{x} = \mathbf{b}$ is solvable is to use elimination on the augmented matrix. If a row of A is completely eliminated, so is the corresponding entry in $\mathbf{b}$. In our example, row 3 of A is completely eliminated:

$$\begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 2 & 4 & 6 & 8 & b_2 \\ 3 & 6 & 8 & 10 & b_3 \end{bmatrix} \rightarrow \cdots \rightarrow \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 0 & 0 & 2 & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_2 - b_1 \end{bmatrix}.$$

If $A\mathbf{x} = \mathbf{b}$ has a solution, then $b_3 - b_2 - b_1 = 0$. For example, we could choose $\mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix}$.

From an earlier lecture, we know that $A\mathbf{x} = \mathbf{b}$ is solvable exactly when $\mathbf{b}$ is in the column space $C(A)$. We have these two conditions on $\mathbf{b}$; in fact they are equivalent.

Complete solution

In order to find all solutions to $A\mathbf{x} = \mathbf{b}$ we first check that the equation is solvable, then find a particular solution. We get the complete solution of the equation by adding the particular solution to all the vectors in the nullspace.

A particular solution

One way to find a particular solution to the equation $A\mathbf{x} = \mathbf{b}$ is to set all free variables to zero, then solve for the pivot variables.

For our example matrix A, we let $x_2 = x_4 = 0$ to get the system of equations:

$$\begin{aligned} x_1 + 2x_3 &= 1 \\ 2x_3 &= 3 \end{aligned}$$

which has the solution $x_3 = 3/2$, $x_1 = -2$. Our particular solution is:

$$\mathbf{x}_p = \begin{bmatrix} -2 \\ 0 \\ 3/2 \\ 0 \end{bmatrix}.$$

Combined with the nullspace

The general solution to $A\mathbf{x} = \mathbf{b}$ is given by $\mathbf{x}_{\text{complete}} = \mathbf{x}_p + \mathbf{x}_n$, where $\mathbf{x}_n$ is a generic vector in the nullspace. To see this, we add $A\mathbf{x}_p = \mathbf{b}$ to $A\mathbf{x}_n = \mathbf{0}$ and get $A\left(\mathbf{x}_p + \mathbf{x}_n\right) = \mathbf{b}$ for every vector $\mathbf{x}_n$ in the nullspace.

Last lecture we learned that the nullspace of A is the collection of all combinations of the special solutions $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix}$. So the complete solution to the equation $A\mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix}$ is:

$$\mathbf{x}_{\text{complete}} = \begin{bmatrix} -2 \\ 0 \\ 3/2 \\ 0 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \\ -2 \\ 1 \end{bmatrix},$$

where c_1 and c_2 are real numbers.

The nullspace of A is a two dimensional subspace of $\mathbb{R}^4$, and the solutions to the equation $A\mathbf{x} = \mathbf{b}$ form a plane parallel to that through $x_p = \begin{bmatrix} -2 \\ 0 \\ 3/2 \\ 0 \end{bmatrix}$.

Rank

The rank of a matrix equals the number of pivots of that matrix. If A is an m by n matrix of rank r, we know $r \leq m$ and $r \leq n$.

Full column rank

If $r = n$, then from the previous lecture we know that the nullspace has dimension $n - r = 0$ and contains only the zero vector. There are no free variables or special solutions.

If $A\mathbf{x} = \mathbf{b}$ has a solution, it is unique; there is either 0 or 1 solution. Examples like this, in which the columns are independent, are common in applications.

We know $r \leq m$, so if $r = n$ the number of columns of the matrix is less than or equal to the number of rows. The row reduced echelon form of the

matrix will look like $R = \begin{bmatrix} I \\ 0 \end{bmatrix}$. For any vector $\mathbf{b}$ in $\mathbb{R}^m$ that's not a linear combination of the columns of A, there is no solution to $A\mathbf{x} = \mathbf{b}$.

Full row rank

If $r = m$, then the reduced matrix $R = \begin{bmatrix} I & F \end{bmatrix}$ has no rows of zeros and so there are no requirements for the entries of $\mathbf{b}$ to satisfy. The equation $A\mathbf{x} = \mathbf{b}$ is solvable for every $\mathbf{b}$. There are $n - r = n - m$ free variables, so there are $n - m$ special solutions to $A\mathbf{x} = \mathbf{0}$.

Full row and column rank

If $r = m = n$ is the number of pivots of A, then A is an invertible square matrix and R is the identity matrix. The nullspace has dimension zero, and $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^m$.

Summary

If R is in row reduced form with pivot columns first (rref), the table below summarizes our results.

	$r = m = n$	$r = n < m$	$r = m < n$	$r < m, r < n$
R	I	$\begin{bmatrix} I \\ 0 \end{bmatrix}$	$\begin{bmatrix} I & F \end{bmatrix}$	$\begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$
# solutions to $A\mathbf{x} = \mathbf{b}$	1	0 or 1	infinitely many	0 or infinitely many

Independence, basis, and dimension

What does it mean for vectors to be independent? How does the idea of independence help us describe subspaces like the nullspace?

Linear independence

Suppose A is an m by n matrix with $m < n$ (so $A\mathbf{x} = \mathbf{b}$ has more unknowns than equations). A has at least one free variable, so there are nonzero solutions to $A\mathbf{x} = \mathbf{0}$. A combination of the columns is zero, so the columns of this A are *dependent*.

We say vectors $\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_n$ are *linearly independent* (or just *independent*) if $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = \mathbf{0}$ only when $c_1, c_2, ..., c_n$ are all 0. When those vectors are the columns of A, the only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

Two vectors are independent if they do not lie on the same line. Three vectors are independent if they do not lie in the same plane. Thinking of $A\mathbf{x}$ as a linear combination of the column vectors of A, we see that the column vectors of A are independent exactly when the nullspace of A contains only the zero vector.

If the columns of A are independent then all columns are pivot columns, the rank of A is n, and there are no free variables. If the columns of A are dependent then the rank of A is less than n and there are free variables.

Spanning a space

Vectors $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k$ *span* a space when the space consists of all combinations of those vectors. For example, the column vectors of A span the column space of A.

If vectors $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_k$ span a space S, then S is the smallest space containing those vectors.

Basis and dimension

A *basis* for a vector space is a sequence of vectors $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_d$ with two properties:

- $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_d$ are independent
- $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_d$ span the vector space.

The basis of a space tells us everything we need to know about that space.

Example: $\mathbb{R}^3$

One basis for $\mathbb{R}^3$ is $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$. These are independent because:

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

is only possible when $c_1 = c_2 = c_3 = 0$. These vectors span $\mathbb{R}^3$.

As discussed at the start of Lecture 10, the vectors $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 3 \\ 8 \end{bmatrix}$ do not form a basis for $\mathbb{R}^3$ because these are the column vectors of a matrix that has two identical rows. The three vectors are not linearly independent.

In general, n vectors in $\mathbb{R}^n$ form a basis if they are the column vectors of an invertible matrix.

Basis for a subspace

The vectors $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 2 \\ 5 \end{bmatrix}$ span a plane in $\mathbb{R}^3$ but they cannot form a basis for $\mathbb{R}^3$. Given a space, every basis for that space has the same number of vectors; that number is the *dimension* of the space. So there are exactly n vectors in every basis for $\mathbb{R}^n$.

Bases of a column space and nullspace

Suppose:

$$A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 1 \end{bmatrix}.$$

By definition, the four column vectors of A span the column space of A. The third and fourth column vectors are dependent on the first and second, and the first two columns are independent. Therefore, the first two column vectors are the pivot columns. They form a basis for the column space $C(A)$. The matrix has rank 2. In fact, for any matrix A we can say:

$$\text{rank}(A) = \text{number of pivot columns of } A = \text{dimension of } C(A).$$

(Note that matrices have a rank but not a dimension. Subspaces have a dimension but not a rank.)

The column vectors of this A are not independent, so the nullspace $N(A)$ contains more than just the zero vector. Because the third column is the sum

of the first two, we know that the vector $\begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ is in the nullspace. Similarly, $\begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$ is also in $N(A)$. These are the two special solutions to $A\mathbf{x} = \mathbf{0}$. We'll see that:

$$\text{dimension of } N(A) = \text{number of free variables} = n - r,$$

so we know that the dimension of $N(A)$ is $4 - 2 = 2$. These two special solutions form a basis for the nullspace.

The four fundamental subspaces

In this lecture we discuss the four fundamental spaces associated with a matrix and the relations between them.

Four subspaces

Any m by n matrix A determines four subspaces (possibly containing only the zero vector):

Column space, $C(A)$

$C(A)$ consists of all combinations of the columns of A and is a vector space in $\mathbb{R}^m$.

Nullspace, $N(A)$

This consists of all solutions $\mathbf{x}$ of the equation $A\mathbf{x} = \mathbf{0}$ and lies in $\mathbb{R}^n$.

Row space, $C(A^T)$

The combinations of the row vectors of A form a subspace of R^n. We equate this with $C(A^T)$, the column space of the transpose of A.

Left nullspace, $N(A^T)$

We call the nullspace of A^T the *left nullspace* of A. This is a subspace of $\mathbb{R}^m$.

Basis and Dimension

Column space

The r pivot columns form a basis for $C(A)$

$$\dim C(A) = r.$$

Nullspace

The special solutions to $A\mathbf{x} = \mathbf{0}$ correspond to free variables and form a basis for $N(A)$. An m by n matrix has $n - r$ free variables:

$$\dim N(A) = n - r.$$

Row space

We could perform row reduction on A^T, but instead we make use of R, the row reduced echelon form of A.

$$A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix} = R$$

Although the column spaces of A and R are different, the row space of R is the same as the row space of A. The rows of R are combinations of the rows of A, and because reduction is reversible the rows of A are combinations of the rows of R.

The first r rows of R are the "echelon" basis for the row space of A:

$$\dim C(A^T) = r.$$

Left nullspace

The matrix A^T has m columns. We just saw that r is the rank of A^T, so the number of free columns of A^T must be $m - r$:

$$\dim N(A^T) = m - r.$$

The left nullspace is the collection of vectors y for which $A^T y = 0$. Equivalently, $y^T A = 0$; here y and 0 are row vectors. We say "left nullspace" because y^T is on the left of A in this equation.

To find a basis for the left nullspace we reduce an augmented version of A:

$$\begin{bmatrix} A_{m\times n} & I_{m\times n} \end{bmatrix} \longrightarrow \begin{bmatrix} R_{m\times n} & E_{m\times n} \end{bmatrix}.$$

From this we get the matrix E for which $EA = R$. (If A is a square, invertible matrix then $E = A^{-1}$.) In our example,

$$EA = \begin{bmatrix} -1 & 2 & 0 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R.$$

The bottom $m - r$ rows of E describe linear dependencies of rows of A, because the bottom $m - r$ rows of R are zero. Here $m - r = 1$ (one zero row in R).

The bottom $m - r$ rows of E satisfy the equation $\mathbf{y}^T A = \mathbf{0}$ and form a basis for the left nullspace of A.

New vector space

The collection of all 3×3 matrices forms a vector space; call it M. We can add matrices and multiply them by scalars and there's a zero matrix (additive identity). If we ignore the fact that we can multiply matrices by each other, they behave just like vectors.

Some subspaces of M include:

- all upper triangular matrices
- all symmetric matrices
- D, all diagonal matrices

D is the intersection of the first two spaces. Its dimension is 3; one basis for D is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

Matrix spaces; rank 1; small world graphs

We've talked a lot about $\mathbb{R}^n$, but we can think about vector spaces made up of any sort of "vectors" that allow addition and scalar multiplication.

New vector spaces

3 by 3 matrices

We were looking at the space M of all 3 by 3 matrices. We identified some subspaces; the symmetric 3 by 3 matrices S, the upper triangular 3 by 3 matrices U, and the intersection D of these two spaces – the space of diagonal 3 by 3 matrices.

The dimension of M is 9; we must choose 9 numbers to specify an element of M. The space M is very similar to $\mathbb{R}^9$. A good choice of basis is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ldots \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The subspace of symmetric matrices S has dimension 6. When choosing an element of S we pick three numbers on the diagonal and three in the upper right, which tell us what must appear in the lower left of the matrix. One basis for S is the collection:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The dimension of U is again 6; we have the same amount of freedom in selecting the entries of an upper triangular matrix as we did in choosing a symmetric matrix. A basis for U is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This happens to be a subset of the basis we chose for M, but there is no basis for S that is a subset of the basis we chose for M.

The subspace $D = S \cap U$ of diagonal 3 by 3 matrices has dimension 3. Because of the way we chose bases for U and S, a good basis for D is the intersection of those bases.

Is $S \cup U$, the set of 3 by 3 matrices which are either symmetric or upper triangular, a subspace of M? No. This is like taking two lines in $\mathbb{R}^2$ and asking if together they form a subspace; we have to fill in between them. If we take all possible sums of elements of S and elements of U we get what we call the *sum* $S + U$. This is a subspace of M. In fact, $S + U = M$. For unions and sums, dimensions follow this rule:

$$\dim S + \dim U = \dim S \cup U + \dim S \cap U.$$

Differential equations

Another example of a vector space that's not $\mathbb{R}^n$ appears in differential equations.

We can think of the solutions y to $\dfrac{d^2y}{dx^2} + y = 0$ as the elements of a nullspace. Some solutions are:

$$y = \cos x, \quad y = \sin x, \quad \text{and} \quad y = e^{ix}.$$

The complete solution is:

$$y = c_1 \cos x + c_2 \sin x,$$

where c_1 and c_2 can be any complex numbers. This solution space is a two dimensional vector space with basis vectors $\cos x$ and $\sin x$. (Even though these don't "look like" vectors, we can build a vector space from them because they can be added and multiplied by a constant.)

Rank 4 matrices

Now let M be the space of 5×17 matrices. The subset of M containing all rank 4 matrices is not a subspace, even if we include the zero matrix, because the sum of two rank 4 matrices may not have rank 4.

In $\mathbb{R}^4$, the set of all vectors $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}$ for which $v_1 + v_2 + v_3 + v_4 = 0$ is a subspace. It contains the zero vector and is closed under addition and scalar multiplication. It is the nullspace of the matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$. Because A has rank 1, the dimension of this nullspace is $n - r = 3$. The subspace has the basis of special solutions:

$$\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The column space of A is $\mathbb{R}^1$. The left nullspace contains only the zero vector, has dimension zero, and its basis is the empty set. The row space of A also has dimension 1.

Rank one matrices

The rank of a matrix is the dimension of its column (or row) space. The matrix

$$A = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 8 & 10 \end{bmatrix}$$

has rank 1 because each of its columns is a multiple of the first column.

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

Every rank 1 matrix A can be written $A = \mathbf{U}\mathbf{V}^T$, where $\mathbf{U}$ and $\mathbf{V}$ are column vectors. We'll use rank 1 matrices as building blocks for more complex matrices.

Small world graphs

In this class, a *graph* G is a collection of nodes joined by edges:

$$G = \{\text{nodes}, \text{edges}\}.$$

A typical graph appears in Figure 1. Another example of a graph is one in

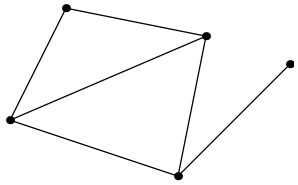

Figure 1: A graph with 5 nodes and 6 edges.

which each node is a person. Two nodes are connected by an edge if the people are friends. We can ask how close two people are to each other in the graph – what's the smallest number of friend to friend connections joining them? The question "what's the farthest distance between two people in the graph?" lies behind phrases like "six degrees of separation" and "it's a small world".

Another graph is the world wide web: its nodes are web sites and its edges are links.

We'll describe graphs in terms of matrices, which will make it easy to answer questions about distances between nodes.

Graphs, networks, incidence matrices

When we use linear algebra to understand physical systems, we often find more structure in the matrices and vectors than appears in the examples we make up in class. There are many applications of linear algebra; for example, chemists might use row reduction to get a clearer picture of what elements go into a complicated reaction. In this lecture we explore the linear algebra associated with electrical networks.

Graphs and networks

A *graph* is a collection of nodes joined by edges; Figure 1 shows one small graph.

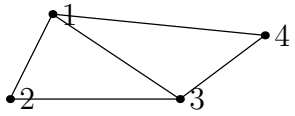

Figure 1: A graph with $n = 4$ nodes and $m = 5$ edges.

We put an arrow on each edge to indicate the positive direction for currents running through the graph.

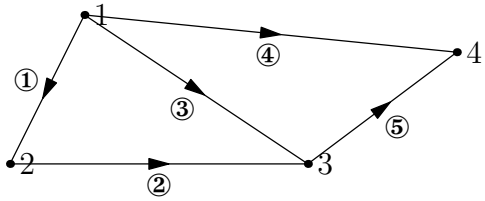

Figure 2: The graph of Figure 1 with a direction on each edge.

Incidence matrices

The *incidence matrix* of this directed graph has one column for each node of the graph and one row for each edge of the graph:

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

If an edge runs from node a to node b, the row corresponding to that edge has -1 in column a and 1 in column b; all other entries in that row are 0. If we were

studying a larger graph we would get a larger matrix but it would be *sparse*; most of the entries in that matrix would be 0. This is one of the ways matrices arising from applications might have extra structure.

Note that nodes 1, 2 and 3 and edges ①, ② and ③ form a loop. The matrix describing just those nodes and edges looks like:

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

Note that the third row is the sum of the first two rows; loops in the graph correspond to linearly dependent rows of the matrix.

To find the nullspace of A, we solve $A\mathbf{x} = \mathbf{0}$:

$$A\mathbf{x} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_3 - x_1 \\ x_4 - x_1 \\ x_4 - x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

If the components x_i of the vector $\mathbf{x}$ describe the electrical potential at the nodes i of the graph, then $A\mathbf{x}$ is a vector describing the *difference* in potential across each edge of the graph. We see $A\mathbf{x} = \mathbf{0}$ when $x_1 = x_2 = x_3 = x_4$, so the nullspace has dimension 1. In terms of an electrical network, the potential difference is zero on each edge if each node has the same potential. We can't tell what that potential is by observing the flow of electricity through the network, but if one node of the network is grounded then its potential is zero. From that we can determine the potential of all other nodes of the graph.

The matrix has 4 columns and a 1 dimensional nullspace, so its rank is 3. The first, second and fourth columns are its pivot columns; these edges connect all the nodes of the graph without forming a loop – a graph with no loops is called a *tree*.

The left nullspace of A consists of the solutions $\mathbf{y}$ to the equation: $A^T\mathbf{y} = \mathbf{0}$. Since A^T has 5 columns and rank 3 we know that the dimension of $N(A^T)$ is $m - r = 2$. Note that 2 is the number of loops in the graph and m is the number of edges. The rank r is $n - 1$, one less than the number of nodes. This gives us # loops = # edges − (# nodes − 1), or:

$$\text{number of nodes} - \text{number of edges} + \text{number of loops} = 1.$$

This is Euler's formula for connected graphs.

Kirchhoff's law

In our example of an electrical network, we started with the potentials x_i of the nodes. The matrix A then told us something about potential differences. An engineer could create a matrix C using Ohm's law and information about

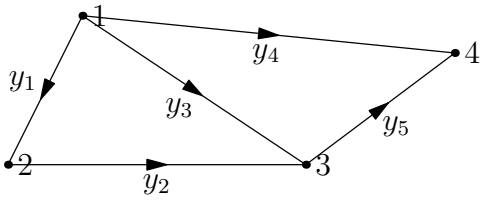

Figure 3: The currents in our graph.

the conductance of the edges and use that matrix to determine the current y_i on each edge. Kirchhoff's Current Law then says that $A^T\mathbf{y} = \mathbf{0}$, where $\mathbf{y}$ is the vector with components y_1, y_2, y_3, y_4, y_5. Vectors in the nullspace of A^T correspond to collections of currents that satisfy Kirchhoff's law.

$\mathbf{x} = x_1, x_2, x_3, x_4$		$A^T\mathbf{y} = 0$
potentials at nodes		Kirchhoff's Current Law
$\mathbf{e} = A\mathbf{x} \downarrow$		$\uparrow A^T\mathbf{y}$
$x_2 - x_1$, etc.	$\mathbf{y} = C\mathbf{e}$	y_1, y_2, y_3, y_4, y_5
potential differences	$\longrightarrow$	currents on edges
	Ohm's Law	

Written out, $A^T\mathbf{y} = \mathbf{0}$ looks like:

$$\begin{bmatrix} -1 & 0 & -1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Multiplying the first row by the column vector $\mathbf{y}$ we get $-y_1 - y_3 - y_4 = 0$. This tells us that the total current flowing out of node 1 is zero – it's a balance equation, or a conservation law. Multiplying the second row by $\mathbf{y}$ tells us $y_1 - y_2 = 0$; the current coming into node 2 is balanced with the current going out. Multiplying the bottom rows, we get $y_2 + y_3 - y_5 = 0$ and $y_4 + y_5 = 0$.

We could use the method of elimination on A^T to find its column space, but we already know the rank. To get a basis for $N(A^T)$ we just need to find two independent vectors in this space. Looking at the equations $y_1 - y_2 = 0$ we might guess $y_1 = y_2 = 1$. Then we could use the conservation laws for node 3 to guess $y_3 = -1$ and $y_5 = 0$. We satisfy the conservation conditions on node 4 with $y_4 = 0$, giving us a basis vector $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$. This vector represents one unit

of current flowing around the loop joining nodes 1, 2 and 3; a multiple of this vector represents a different amount of current around the same loop.

We find a second basis vector for $N(A^T)$ by looking at the loop formed by nodes 1, 3 and 4: $\begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$. The vector $\begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$ that represents a current around the outer loop is also in the nullspace, but it is the sum of the first two vectors we found.

We've almost completely covered the mathematics of simple circuits. More complex circuits might have batteries in the edges, or current sources between nodes. Adding current sources changes the $A^T\mathbf{y} = \mathbf{0}$ in Kirchhoff's current law to $A^T\mathbf{y} = \mathbf{f}$. Combining the equations $\mathbf{e} = A\mathbf{x}$, $\mathbf{y} = C\mathbf{e}$ and $A^T\mathbf{y} = \mathbf{f}$ gives us:

$$A^TCA\mathbf{x} = \mathbf{f}.$$

Orthogonal vectors and subspaces

In this lecture we learn what it means for vectors, bases and subspaces to be *orthogonal*. The symbol for this is $\perp$.

The "big picture" of this course is that the row space of a matrix' is orthogonal to its nullspace, and its column space is orthogonal to its left nullspace.

row space	column space
dimension r	dimension r
$\perp$	$\perp$
nullspace	left nullspace $N(A^T)$
dimension $n-r$	dimension $m-r$

Orthogonal vectors

Orthogonal is just another word for *perpendicular*. Two vectors are *orthogonal* if the angle between them is 90 degrees. If two vectors are orthogonal, they form a right triangle whose hypotenuse is the sum of the vectors. Thus, we can use the Pythagorean theorem to prove that *the dot product* $\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$ *is zero* exactly when $\mathbf{x}$ and $\mathbf{y}$ are orthogonal. (The length squared $||\mathbf{x}||^2$ equals $\mathbf{x}^T\mathbf{x}$.)

Note that all vectors are orthogonal to the zero vector.

Orthogonal subspaces

Subspace S is *orthogonal* to subspace T means: every vector in S is orthogonal to every vector in T. The blackboard is not orthogonal to the floor; two vectors in the line where the blackboard meets the floor aren't orthogonal to each other.

In the plane, the space containing only the zero vector and any line through the origin are orthogonal subspaces. A line through the origin and the whole plane are never orthogonal subspaces. Two lines through the origin are orthogonal subspaces if they meet at right angles.

Nullspace is perpendicular to row space

The row space of a matrix is orthogonal to the nullspace, because $A\mathbf{x} = \mathbf{0}$ means the dot product of $\mathbf{x}$ with each row of A is 0. But then the product of $\mathbf{x}$ with any combination of rows of A must be 0.

The column space is orthogonal to the left nullspace of A because the row space of A^T is perpendicular to the nullspace of A^T.

In some sense, the row space and the nullspace of a matrix subdivide $\mathbb{R}^n$ into two perpendicular subspaces. For $A = \begin{bmatrix} 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}$, the row space has dimension 1 and basis $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$ and the nullspace has dimension 2 and is the

plane through the origin perpendicular to the vector $\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$.

Not only is the nullspace orthogonal to the row space, their dimensions add up to the dimension of the whole space. We say that the nullspace and the row space are *orthogonal complements* in $\mathbb{R}^n$. The nullspace contains all the vectors that are perpendicular to the row space, and vice versa.

We could say that this is part two of the fundamental theorem of linear algebra. Part one gives the dimensions of the four subspaces, part two says those subspaces come in orthogonal pairs, and part three will be about orthogonal bases for these subspaces.

$N(A^TA) = N(A)$

Due to measurement error, $A\mathbf{x} = \mathbf{b}$ is often unsolvable if $m > n$. Our next challenge is to find the best possible solution in this case. The matrix A^TA plays a key role in this effort: the central equation is $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$.

We know that A^TA is square ($n \times n$) and symmetric. When is it invertible?

Suppose $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 5 \end{bmatrix}$. Then:

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 3 & 8 \\ 8 & 30 \end{bmatrix}$$

is invertible. A^TA is not always invertible. In fact:

$$\begin{aligned} N(A^TA) &= N(A) \\ \text{rank of } A^TA &= \text{rank of } A. \end{aligned}$$

We conclude that A^TA is invertible exactly when A has independent columns.

Projections onto subspaces

Projections

If we have a vector $\mathbf{b}$ and a line determined by a vector $\mathbf{a}$, how do we find the point on the line that is closest to $\mathbf{b}$?

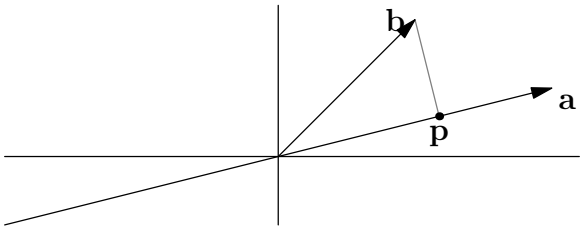

Figure 1: The point closest to $\mathbf{b}$ on the line determined by $\mathbf{a}$.

We can see from Figure 1 that this closest point $\mathbf{p}$ is at the intersection formed by a line through $\mathbf{b}$ that is orthogonal to $\mathbf{a}$. If we think of $\mathbf{p}$ as an approximation of $\mathbf{b}$, then the length of $\mathbf{e} = \mathbf{b} - \mathbf{p}$ is the error in that approximation.

We could try to find $\mathbf{p}$ using trigonometry or calculus, but it's easier to use linear algebra. Since $\mathbf{p}$ lies on the line through $\mathbf{a}$, we know $\mathbf{p} = x\mathbf{a}$ for some number x. We also know that $\mathbf{a}$ is perpendicular to $\mathbf{e} = \mathbf{b} - \mathbf{x}\mathbf{a}$:

$$\begin{aligned} \mathbf{a}^T(\mathbf{b} - x\mathbf{a}) &= 0 \\ x\mathbf{a}^T\mathbf{a} &= \mathbf{a}^T\mathbf{b} \\ x &= \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}, \end{aligned}$$

and $\mathbf{p} = \mathbf{a}x = \mathbf{a}\dfrac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$. Doubling $\mathbf{b}$ doubles $\mathbf{p}$. Doubling $\mathbf{a}$ does not affect $\mathbf{p}$.

Projection matrix

We'd like to write this projection in terms of a *projection matrix* P: $\mathbf{p} = P\mathbf{b}$.

$$\mathbf{p} = \mathbf{x}\mathbf{a} = \frac{\mathbf{a}\mathbf{a}^T\mathbf{a}}{\mathbf{a}^T\mathbf{a}},$$

so the matrix is:

$$P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}.$$

Note that $\mathbf{a}\mathbf{a}^T$ is a three by three matrix, not a number; matrix multiplication is not commutative.

The column space of P is spanned by $\mathbf{a}$ because for any $\mathbf{b}$, $P\mathbf{b}$ lies on the line determined by $\mathbf{a}$. The rank of P is 1. P is symmetric. $P^2\mathbf{b} = P\mathbf{b}$ because

the projection of a vector already on the line through $\mathbf{a}$ is just that vector. In general, projection matrices have the properties:

$$P^T = P \qquad \text{and} \qquad P^2 = P.$$

Why project?

As we know, the equation $A\mathbf{x} = \mathbf{b}$ may have no solution. The vector $A\mathbf{x}$ is always in the column space of A, and $\mathbf{b}$ is unlikely to be in the column space. So, we project $\mathbf{b}$ onto a vector $\mathbf{p}$ in the column space of A and solve $A\hat{\mathbf{x}} = \mathbf{p}$.

Projection in higher dimensions

In $\mathbb{R}^3$, how do we project a vector $\mathbf{b}$ onto the closest point $\mathbf{p}$ in a plane?

If $\mathbf{a}_1$ and $\mathbf{a}_2$ form a basis for the plane, then that plane is the column space of the matrix $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$.

We know that $\mathbf{p} = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2 = A\hat{\mathbf{x}}$. We want to find $\hat{\mathbf{x}}$. There are many ways to show that $\mathbf{e} = \mathbf{b} - \mathbf{p} = \mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to the plane we're projecting onto, after which we can use the fact that $\mathbf{e}$ is perpendicular to $\mathbf{a}_1$ and $\mathbf{a}_2$:

$$\mathbf{a}_1^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \quad \text{and} \quad \mathbf{a}_2^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0.$$

In matrix form, $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$. When we were projecting onto a line, A only had one column and so this equation looked like: $a^T(\mathbf{b} - x\mathbf{a}) = \mathbf{0}$.

Note that $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$ is in the nullspace of A^T and so is in the left nullspace of A. We know that everything in the left nullspace of A is perpendicular to the column space of A, so this is another confirmation that our calculations are correct.

We can rewrite the equation $A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$ as:

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

When projecting onto a line, A^TA was just a number; now it is a square matrix. So instead of dividing by $\mathbf{a}^T\mathbf{a}$ we now have to multiply by $(A^TA)^{-1}$

In n dimensions,

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ \mathbf{p} = A\hat{\mathbf{x}} &= A(A^TA)^{-1}A^T\mathbf{b} \\ P &= A(A^TA)^{-1}A^T. \end{aligned}$$

It's tempting to try to simplify these expressions, but if A isn't a square matrix we can't say that $(A^TA)^{-1} = A^{-1}(A^T)^{-1}$. If A does happen to be a square, invertible matrix then its column space is the whole space and contains $\mathbf{b}$. In this case P is the identity, as we find when we simplify. It is still true that:

$$P^T = P \qquad \text{and} \qquad P^2 = P.$$

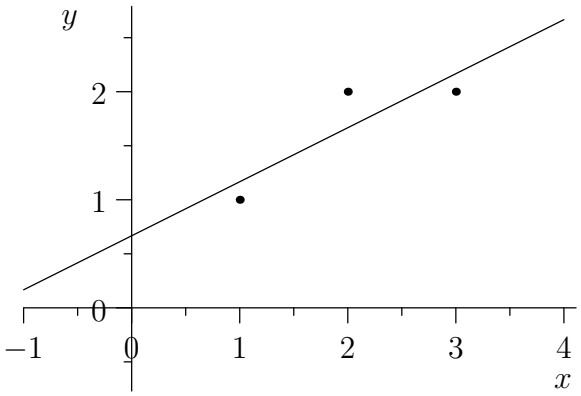

Figure 2: Three points and a line close to them.

Least Squares

Suppose we're given a collection of data points (t, b):

$$\{(1,1),(2,2),(3,2)\}$$

and we want to find the closest line $b = C + Dt$ to that collection. If the line went through all three points, we'd have:

$$\begin{aligned} C + D &= 1 \\ C + 2D &= 2 \\ C + 3D &= 2, \end{aligned}$$

which is equivalent to:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \quad \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{\mathbf{x}} \quad = \quad \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{\mathbf{b}} .$$

In our example the line does not go through all three points, so this equation is not solvable. Instead we'll solve:

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

Projection matrices and least squares

Projections

Last lecture, we learned that $P = A(A^TA)^{-1}A^T$ is the matrix that projects a vector $\mathbf{b}$ onto the space spanned by the columns of A. If $\mathbf{b}$ is perpendicular to the column space, then it's in the left nullspace $N(A^T)$ of A and $P\mathbf{b} = \mathbf{0}$. If $\mathbf{b}$ is in the column space then $\mathbf{b} = A\mathbf{x}$ for some $\mathbf{x}$, and $P\mathbf{b} = \mathbf{b}$.

A typical vector will have a component $\mathbf{p}$ in the column space and a component $\mathbf{e}$ perpendicular to the column space (in the left nullspace); its projection is just the component in the column space.

The matrix projecting $\mathbf{b}$ onto $N(A^T)$ is $I - P$:

$$\begin{aligned} \mathbf{e} &= \mathbf{b} - \mathbf{p} \\ \mathbf{e} &= (I - P)\mathbf{b}. \end{aligned}$$

Naturally, $I - P$ has all the properties of a projection matrix.

Least squares

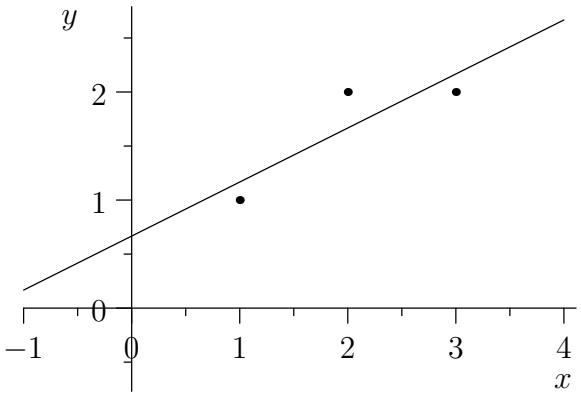

Figure 1: Three points and a line close to them.

We want to find the closest line $b = C + Dt$ to the points $(1, 1)$, $(2, 2)$, and $(3, 2)$. The process we're going to use is called *linear regression*; this technique is most useful if none of the data points are *outliers*.

By "closest" line we mean one that minimizes the error represented by the distance from the points to the line. We measure that error by adding up the squares of these distances. In other words, we want to minimize $||A\mathbf{x} - \mathbf{b}||^2 = ||\mathbf{e}||^2$.

If the line went through all three points, we'd have:

$$\begin{aligned} C + D &= 1 \\ C + 2D &= 2 \\ C + 3D &= 2, \end{aligned}$$

but this system is unsolvable. It's equivalent to $A\mathbf{x} = \mathbf{b}$, where:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} C \\ D \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

There are two ways of viewing this. In the space of the line we're trying to find, e_1, e_2 and e_3 are the vertical distances from the data points to the line. The components p_1, p_2 and p_3 are the values of $C + Dt$ near each data point; $\mathbf{p} \approx \mathbf{b}$.

In the other view we have a vector $\mathbf{b}$ in $\mathbb{R}^3$, its projection $\mathbf{p}$ onto the column space of A, and its projection $\mathbf{e}$ onto $N(A^T)$.

We will now find $\hat{\mathbf{x}} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$ and $\mathbf{p}$. We know:

$$\begin{aligned} A^T A\hat{\mathbf{x}} &= A^T\mathbf{b} \\ \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} &= \begin{bmatrix} 5 \\ 11 \end{bmatrix}. \end{aligned}$$

From this we get the *normal equations*:

$$\begin{aligned} 3\hat{C} + 6\hat{D} &= 5 \\ 6\hat{C} + 14\hat{D} &= 11. \end{aligned}$$

We solve these to find $\hat{D} = 1/2$ and $\hat{C} = 2/3$.

We could also have used calculus to find the minimum of the following function of two variables:

$$e_1^2 + e_2^2 + e_3^2 = (C + D - 1)^2 + (C + 2D - 2)^2 + (C + 3D - 2)^2.$$

Either way, we end up solving a system of linear equations to find that the closest line to our points is $b = \frac{2}{3} + \frac{1}{2}t$.

This gives us:

i	p_i	e_i
1	7/6	−1/6
2	5/3	1/3
3	13/6	−1/6

or $\mathbf{p} = \begin{bmatrix} 7/6 \\ 5/3 \\ 13/6 \end{bmatrix}$ and $\mathbf{e} = \begin{bmatrix} -1/6 \\ 2/6 \\ -1/6 \end{bmatrix}$. Note that $\mathbf{p}$ and $\mathbf{e}$ are orthogonal, and also that $\mathbf{e}$ is perpendicular to the columns of A.

The matrix $A^T A$

We've been assuming that the matrix $A^T A$ is invertible. Is this justified?

If A has independent columns, then $A^T A$ is invertible.

To prove this we assume that $A^T A\mathbf{x} = \mathbf{0}$, then show that it must be true that $\mathbf{x} = \mathbf{0}$:

$$\begin{aligned} A^T A\mathbf{x} &= \mathbf{0} \\ \mathbf{x}^T A^T A\mathbf{x} &= \mathbf{x}^T\mathbf{0} \\ (A\mathbf{x})^T(A\mathbf{x}) &= \mathbf{0} \\ A\mathbf{x} &= \mathbf{0}. \end{aligned}$$

Since A has independent columns, $A\mathbf{x} = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$.

As long as the columns of A are independent, we can use linear regression to find approximate solutions to unsolvable systems of linear equations. The columns of A are guaranteed to be independent if they are *orthonormal*, i.e. if they are perpendicular unit vectors like $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, or like $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ and $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$.

Orthogonal matrices and Gram-Schmidt

In this lecture we finish introducing orthogonality. Using an orthonormal basis or a matrix with orthonormal columns makes calculations much easier. The Gram-Schmidt process starts with any basis and produces an orthonormal basis that spans the same space as the original basis.

Orthonormal vectors

The vectors $\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_n$ are *orthonormal* if:

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

In other words, they all have (normal) length 1 and are perpendicular (ortho) to each other. Orthonormal vectors are always independent.

Orthonormal matrix

If the columns of $Q = \begin{bmatrix} \mathbf{q}_1 & ... & \mathbf{q}_n \end{bmatrix}$ are orthonormal, then $Q^T Q = I$ is the identity.

Matrices with orthonormal columns are a new class of important matrices to add to those on our list: triangular, diagonal, permutation, symmetric, reduced row echelon, and projection matrices. We'll call them "orthonormal matrices".

A square orthonormal matrix Q is called an *orthogonal matrix*. If Q is square, then $Q^T Q = I$ tells us that $Q^T = Q^{-1}$.

For example, if $Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ then $Q^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Both Q and Q^T are orthogonal matrices, and their product is the identity.

The matrix $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal. The matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is not, but we can adjust that matrix to get the orthogonal matrix $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

We can use the same tactic to find some larger orthogonal matrices called *Hadamard matrices*:

$$Q = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

An example of a rectangular matrix with orthonormal columns is:

$$Q = \frac{1}{3} \begin{bmatrix} 1 & -2 \\ 2 & -1 \\ 2 & 2 \end{bmatrix}.$$

We can extend this to a (square) orthogonal matrix:

$$\frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}.$$

These examples are particularly nice because they don't include complicated square roots.

Orthonormal columns are good

Suppose Q has orthonormal columns. The matrix that projects onto the column space of Q is:

$$P = Q^T(Q^TQ)^{-1}Q^T.$$

If the columns of Q are orthonormal, then $Q^TQ = I$ and $P = QQ^T$. If Q is square, then $P = I$ because the columns of Q span the entire space.

Many equations become trivial when using a matrix with orthonormal columns. If our basis is orthonormal, the projection component $\hat{x}_i$ is just $\mathbf{q}_i^T\mathbf{b}$ because $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ becomes $\hat{\mathbf{x}} = Q^T\mathbf{b}$.

Gram-Schmidt

With elimination, our goal was "make the matrix triangular". Now our goal is "make the matrix orthonormal".

We start with two independent vectors $\mathbf{a}$ and $\mathbf{b}$ and want to find orthonormal vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ that span the same plane. We start by finding orthogonal vectors $\mathbf{A}$ and $\mathbf{B}$ that span the same space as $\mathbf{a}$ and $\mathbf{b}$. Then the unit vectors $\mathbf{q}_1 = \frac{\mathbf{A}}{||\mathbf{A}||}$ and $\mathbf{q}_2 = \frac{\mathbf{B}}{||\mathbf{B}||}$ form the desired orthonormal basis.

Let $\mathbf{A} = \mathbf{a}$. We get a vector orthogonal to $\mathbf{A}$ in the space spanned by $\mathbf{a}$ and $\mathbf{b}$ by projecting $\mathbf{b}$ onto $\mathbf{a}$ and letting $\mathbf{B} = \mathbf{b} - \mathbf{p}$. ($\mathbf{B}$ is what we previously called $\mathbf{e}$.)

$$\mathbf{B} = \mathbf{b} - \frac{\mathbf{A}^T\mathbf{b}}{\mathbf{A}^T\mathbf{A}}\mathbf{A}.$$

If we multiply both sides of this equation by $\mathbf{A}^T$, we see that $\mathbf{A}^T\mathbf{B} = 0$.

What if we had started with three independent vectors, $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$? Then we'd find a vector $\mathbf{C}$ orthogonal to both $\mathbf{A}$ and $\mathbf{B}$ by subtracting from $\mathbf{c}$ its components in the $\mathbf{A}$ and $\mathbf{B}$ directions:

$$\mathbf{C} = \mathbf{c} - \frac{\mathbf{A}^T\mathbf{c}}{\mathbf{A}^T\mathbf{A}}\mathbf{A} - \frac{\mathbf{B}^T\mathbf{c}}{\mathbf{B}^T\mathbf{B}}\mathbf{B}.$$

For example, suppose $\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$. Then $\mathbf{A} = \mathbf{a}$ and:

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \frac{\mathbf{A}^T\mathbf{b}}{\mathbf{A}^T\mathbf{A}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}. \end{aligned}$$

Normalizing, we get:

$$Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix}.$$

The column space of Q is the plane spanned by $\mathbf{a}$ and $\mathbf{b}$.

When we studied elimination, we wrote the process in terms of matrices and found $A = LU$. A similar equation $A = QR$ relates our starting matrix A to the result Q of the Gram-Schmidt process. Where L was lower triangular, R is upper triangular.

Suppose $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$. Then:

$$\begin{matrix} A \\ \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} \end{matrix} = \begin{matrix} Q \\ \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 \end{bmatrix} \end{matrix} \begin{matrix} R \\ \begin{bmatrix} \mathbf{a}_1^T\mathbf{q}_1 & \mathbf{a}_2^T\mathbf{q}_1 \\ \mathbf{a}_1^T\mathbf{q}_2 & \mathbf{a}_2^T\mathbf{q}_2 \end{bmatrix} \end{matrix}.$$

If R is upper triangular, then it should be true that $\mathbf{a}_1^T\mathbf{q}_2 = 0$. This must be true because we chose $\mathbf{q}_1$ to be a unit vector in the direction of $\mathbf{a}_1$. All the later $\mathbf{q}_i$ were chosen to be perpendicular to the earlier ones.

Notice that $R = Q^TA$. This makes sense; $Q^TQ = I$.

Properties of determinants

Determinants

Now halfway through the course, we leave behind rectangular matrices and focus on square ones. Our next big topics are determinants and eigenvalues.

The *determinant* is a number associated with any square matrix; we'll write it as $\det A$ or $|A|$. The determinant encodes a lot of information about the matrix; the matrix is invertible exactly when the determinant is non-zero.

Properties

Rather than start with a big formula, we'll list the properties of the determinant. We already know that $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$; these properties will give us a formula for the determinant of square matrices of all sizes.

1. $\det I = 1$
2. If you exchange two rows of a matrix, you reverse the sign of its determinant from positive to negative or from negative to positive.
3. (a) If we multiply one row of a matrix by t, the determinant is multiplied by t: $\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

   (b) The determinant behaves like a linear function on the rows of the matrix:
   $$\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}.$$

Property 1 tells us that $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$. Property 2 tells us that $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$. The determinant of a permutation matrix P is 1 or -1 depending on whether P exchanges an even or odd number of rows.

From these three properties we can deduce many others:

4. If two rows of a matrix are equal, its determinant is zero.

   This is because of property 2, the exchange rule. On the one hand, exchanging the two identical rows does not change the determinant. On the other hand, exchanging the two rows changes the sign of the determinant. Therefore the determinant must be 0.

5. If $i \neq j$, subtracting t times row i from row j doesn't change the determinant.

In two dimensions, this argument looks like:

$$\begin{vmatrix} a & b \\ c-ta & d-tb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - \begin{vmatrix} a & b \\ ta & tb \end{vmatrix} \quad \text{property 3(b)}$$

$$= \begin{vmatrix} a & b \\ c & d \end{vmatrix} - t \begin{vmatrix} a & b \\ a & b \end{vmatrix} \quad \text{property 3(a)}$$

$$= \begin{vmatrix} a & b \\ c & d \end{vmatrix} \quad \text{property 4.}$$

The proof for higher dimensional matrices is similar.

6. If A has a row that is all zeros, then $\det A = 0$.

   We get this from property 3 (a) by letting $t = 0$.

7. The determinant of a triangular matrix is the product of the diagonal entries (pivots) $d_1, d_2, ..., d_n$.

   Property 5 tells us that the determinant of the triangular matrix won't change if we use elimination to convert it to a diagonal matrix with the entries d_i on its diagonal. Then property 3 (a) tells us that the determinant of this diagonal matrix is the product $d_1 d_2 \cdots d_n$ times the determinant of the identity matrix. Property 1 completes the argument.

   Note that we cannot use elimination to get a diagonal matrix if one of the d_i is zero. In that case elimination will give us a row of zeros and property 6 gives us the conclusion we want.

8. $\det A = 0$ exactly when A is singular.

   If A is singular, then we can use elimination to get a row of zeros, and property 6 tells us that the determinant is zero.

   If A is not singular, then elimination produces a full set of pivots $d_1, d_2, ..., d_n$ and the determinant is $d_1 d_2 \cdots d_n \neq 0$ (with minus signs from row exchanges).

We now have a very practical formula for the determinant of a non-singular matrix. In fact, the way computers find the determinants of large matrices is to first perform elimination (keeping track of whether the number of row exchanges is odd or even) and then multiply the pivots:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \longrightarrow \begin{bmatrix} a & b \\ 0 & d - \frac{c}{a}b \end{bmatrix}, \text{ if } a \neq 0, \text{ so}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a(d - \frac{c}{a}b) = ad - bc.$$

9. $\det AB = (\det A)(\det B)$

   This is very useful. Although the determinant of a sum does not equal the sum of the determinants, it is true that the determinant of a product equals the product of the determinants.

For example:

$$\det A^{-1} = \frac{1}{\det A},$$

because $A^{-1}A = 1$. (Note that if A is singular then A^{-1} does not exist and $\det A^{-1}$ is undefined.) Also, $\det A^2 = (\det A)^2$ and $\det 2A = 2^n \det A$ (applying property 3 to each row of the matrix). This reminds us of volume – if we double the length, width and height of a three dimensional box, we increase its volume by a multiple of $2^3 = 8$.

10. $\det A^T = \det A$

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

This lets us translate properties (2, 3, 4, 5, 6) involving rows into statements about columns. For instance, if a column of a matrix is all zeros then the determinant of that matrix is zero.

To see why $|A^T| = |A|$, use elimination to write $A = LU$. The statement becomes $|U^T L^T| = |LU|$. Rule 9 then tells us $|U^T||L^T| = |L||U|$.

Matrix L is a lower triangular matrix with 1's on the diagonal, so rule 5 tells us that $|L| = |L^T| = 1$. Because U is upper triangular, rule 5 tells us that $|U| = |U^T|$. Therefore $|U^T||L^T| = |L||U|$ and $|A^T| = |A|$.

We have one loose end to worry about. Rule 2 told us that a row exchange changes the sign of the determinant. If it's possible to do seven row exchanges and get the same matrix you would by doing ten row exchanges, then we could prove that the determinant equals its negative. To complete the proof that the determinant is well defined by properties 1, 2 and 3 we'd need to show that the result of an odd number of row exchanges (odd permutation) can never be the same as the result of an even number of row exchanges (even permutation).

Determinant formulas and cofactors

Now that we know the properties of the determinant, it's time to learn some (rather messy) formulas for computing it.

Formula for the determinant

We know that the determinant has the following three properties:

1. $\det I = 1$
2. Exchanging rows reverses the sign of the determinant.
3. The determinant is linear in each row separately.

Last class we listed seven consequences of these properties. We can use these ten properties to find a formula for the determinant of a 2 by 2 matrix:

$$\begin{aligned}
\begin{vmatrix} a & b \\ c & d \end{vmatrix} &= \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} \\
&= \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} \\
&= 0 + ad + (-cb) + 0 \\
&= ad - bc.
\end{aligned}$$

By applying property 3 to separate the individual entries of each row we could get a formula for any other square matrix. However, for a 3 by 3 matrix we'll have to add the determinants of twenty seven different matrices! Many of those determinants are zero. The non-zero pieces are:

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix} \\
&\quad + \begin{vmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix} \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{33} - a_{12}a_{21}a_{33} \\
&\quad + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.
\end{aligned}$$

Each of the non-zero pieces has one entry from each row in each column, as in a permutation matrix. Since the determinant of a permutation matrix is either 1 or -1, we can again use property 3 to find the determinants of each of these summands and obtain our formula.

One way to remember this formula is that the positive terms are products of entries going down and to the right in our original matrix, and the negative terms are products going down and to the left. This rule of thumb doesn't work for matrices larger than 3 by 3.

The number of parts with non-zero determinants was 2 in the 2 by 2 case, 6 in the 3 by 3 case, and will be $24 = 4!$ in the 4 by 4 case. This is because there are n ways to choose an element from the first row (i.e. a value for α), after which there are only $n - 1$ ways to choose an element from the second row that avoids a zero determinant. Then there are $n - 2$ choices from the third row, $n - 3$ from the fourth, and so on.

The big formula for computing the determinant of any square matrix is:

$$\det A = \sum_{n! \text{ terms}} \pm a_{1\alpha} a_{2\beta} a_{3\gamma} ... a_{n\omega}$$

where $(\alpha, \beta, \gamma, ...\omega)$ is some permutation of $(1, 2, 3, ..., n)$. If we test this on the identity matrix, we find that all the terms are zero except the one corresponding to the trivial permutation $\alpha = 1$, $\beta = 2$, ..., $\omega = n$. This agrees with the first property: $\det I = 1$. It's possible to check all the other properties as well, but we won't do that here.

Applying the method of elimination and multiplying the diagonal entries of the result (the pivots) is another good way to find the determinant of a matrix.

Example

In a matrix with many zero entries, many terms in the formula are zero. We can compute the determinant of:

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

by choosing a non-zero entry from each row and column, multiplying those entries, giving the product the appropriate sign, then adding the results.

The permutation corresponding to the diagonal running from a_{14} to a_{41} is $(4, 3, 2, 1)$. This contributes 1 to the determinant of the matrix; the contribution is positive because it takes two row exchanges to convert the permutation $(4, 3, 2, 1)$ to the identity $(1, 2, 3, 4)$.

Another non-zero term of $\sum \pm a_{1\alpha} a_{2\beta} a_{3\gamma} a_{4\omega}$ comes from the permutation $(3, 2, 1, 4)$. This contributes -1 to the sum, because one exchange (of the first and third rows) leads to the identity.

These are the only two non-zero terms in the sum, so the determinant is 0. We can confirm this by noting that row 1 minus row 2 plus row 3 minus row 4 equals zero.

Cofactor formula

The cofactor formula rewrites the big formula for the determinant of an n by n matrix in terms of the determinants of smaller matrices.

In the 3×3 case, the formula looks like:

$$\begin{aligned}\det A &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(-a_{21}a_{33} + a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{vmatrix}\end{aligned}$$

This comes from grouping all the multiples of a_{ij} in the big formula. Each element is multiplied by the *cofactors* in the parentheses following it. Note that each cofactor is (plus or minus) the determinant of a two by two matrix. That determinant is made up of products of elements in the rows and columns NOT containing a_{1j}.

In general, the cofactor C_{ij} of a_{ij} can be found by looking at all the terms in the big formula that contain a_{ij}. C_{ij} equals $(-1)^{i+j}$ times the determinant of the $n-1$ by $n-1$ square matrix obtained by removing row i and column j. (C_{ij} is positive if $i+j$ is even and negative if $i+j$ is odd.)

For $n \times n$ matrices, the cofactor formula is:

$$\boxed{\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}.}$$

Applying this to a 2×2 matrix gives us:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad + b(-c).$$

Tridiagonal matrix

A *tridiagonal matrix* is one for which the only non-zero entries lie on or adjacent to the diagonal. For example, the 4×4 tridiagonal matrix of 1's is:

$$A_4 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

What is the determinant of an $n \times n$ tridiagonal matrix of 1's?

$$|A_1| = 1, |A_2| = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0, |A_3| = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -1$$

$$|A_4| = 1\begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{vmatrix} = |A_3| - 1|A_2| = -1$$

In fact, $|A_n| = |A_{n-1}| - |A_{n-2}|$. We get a sequence which repeats every six terms:

$$|A_1| = 1, |A_2| = 0, |A_3| = -1, |A_4| = -1, |A_5| = 0, |A_6| = 1, |A_7| = 1.$$

Cramer's rule, inverse matrix, and volume

We know a formula for and some properties of the determinant. Now we see how the determinant can be used.

Formula for A^{-1}

We know:

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

Can we get a formula for the inverse of a 3 by 3 or n by n matrix? We expect that $\frac{1}{\det A}$ will be involved, as it is in the 2 by 2 example, and by looking at the cofactor matrix $\begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$ we might guess that cofactors will be involved.

In fact:

$$A^{-1} = \frac{1}{\det A}C^T$$

where C is the matrix of cofactors – please notice the transpose! Cofactors of row one of A go into column 1 of A^{-1}, and then we divide by the determinant.

The determinant of A involves products with n terms and the cofactor matrix involves products of $n-1$ terms. A and $\frac{1}{\det A}C^T$ might cancel each other. This is much easier to see from our formula for the determinant than when using Gauss-Jordan elimination.

To more formally verify the formula, we'll check that $AC^T = (\det A)I$.

$$AC^T = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix}.$$

The entry in the first row and first column of the product matrix is:

$$\sum_{j=1}^{n} a_{1j}C_{j1} = \det A.$$

(This is just the cofactor formula for the determinant.) This happens for every entry on the diagonal of AC^T.

To finish proving that $AC^T = (\det A)I$, we just need to check that the off-diagonal entries of AC^T are zero. In the two by two case, multiplying the entries in row 1 of A by the entries in column 2 of C^T gives $a(-b)+b(a) = 0$. This is the determinant of $A_s = \begin{bmatrix} a & b \\ a & b \end{bmatrix}$. In higher dimensions, the product of the first row of A and the last column of C^T equals the determinant of a matrix whose first and last rows are identical. This happens with all the off diagonal matrices, which confirms that $A^{-1} = \frac{1}{\det A}C^T$.

This formula helps us answer questions about how the inverse changes when the matrix changes.

Cramer's Rule for $\mathbf{x} = A^{-1}\mathbf{b}$

We know that if $A\mathbf{x} = \mathbf{b}$ and A is nonsingular, then $\mathbf{x} = A^{-1}\mathbf{b}$. Applying the formula $A^{-1} = C^T/\det A$ gives us:

$$\mathbf{x} = \frac{1}{\det A} C^T \mathbf{b}.$$

Cramer's rule gives us another way of looking at this equation. To derive this rule we break $\mathbf{x}$ down into its components. Because the i'th component of $C^T\mathbf{b}$ is a sum of cofactors times some number, it is the determinant of some matrix B_j.

$$x_j = \frac{\det B_j}{\det A},$$

where B_j is the matrix created by starting with A and then replacing column j with $\mathbf{b}$, so:

$$B_1 = \begin{bmatrix} & \text{last n-1} & \\ \mathbf{b} & \text{columns} & \\ & \text{of } A & \end{bmatrix} \quad \text{and}$$

$$B_n = \begin{bmatrix} & \text{first n-1} & \\ & \text{columns} & \mathbf{b} \\ & \text{of } A & \end{bmatrix}.$$

This agrees with our formula $x_1 = \frac{\det B_1}{\det A}$. When taking the determinant of B_1 we get a sum whose first term is b_1 times the cofactor C_{11} of A.

Computing inverses using Cramer's rule is usually less efficient than using elimination.

$|\det A| =$ volume of box

Claim: $|\det A|$ is the volume of the box (*parallelepiped*) whose edges are the column vectors of A. (We could equally well use the row vectors, forming a different box with the same volume.)

If $A = I$, then the box is a unit cube and its volume is 1. Because this agrees with our claim, we can conclude that the volume obeys determinant property 1.

If $A = Q$ is an orthogonal matrix then the box is a unit cube in a different orientation with volume $1 = |\det Q|$. (Because Q is an orthogonal matrix, $Q^TQ = I$ and so $\det Q = \pm 1$.)

Swapping two columns of A does not change the volume of the box or (remembering that $\det A = \det A^T$) the absolute value of the determinant (property 2). If we show that the volume of the box also obeys property 3 we'll have proven $|\det A|$ equals the volume of the box.

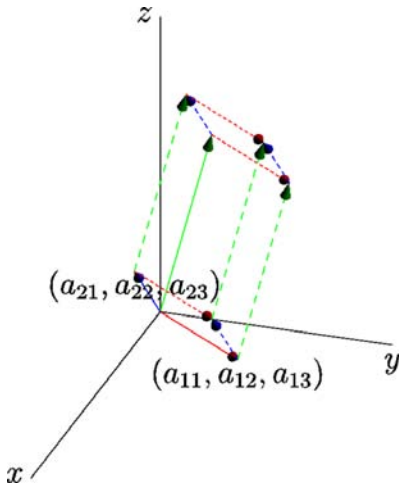

Figure 1: The box whose edges are the column vectors of A.

If we double the length of one column of A, we double the volume of the box formed by its columns. Volume satisfies property 3(a).

Property 3(b) says that the determinant is linear in the rows of the matrix:

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

Figure 2 illustrates why this should be true.

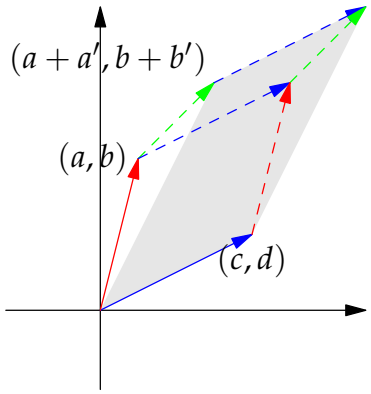

Figure 2: Volume obeys property 3(b).

Although it's not needed for our proof, we can also see that determinants obey property 4. If two edges of a box are equal, the box flattens out and has no volume.

Important note: If you know the coordinates for the corners of a box, then computing the volume of the box is as easy as calculating a determinant. In particular, the area of a parallelogram with edges $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ is $ad - bc$. The area of a triangle with edges $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ is half the area of that parallelogram, or $\frac{1}{2}(ad - bc)$. The area of a triangle with vertices at (x_1, y_1), (x_2, y_2) and (x_3, y_3) is:

$$\frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

Eigenvalues and eigenvectors

The subject of eigenvalues and eigenvectors will take up most of the rest of the course. We will again be working with square matrices. Eigenvalues are special numbers associated with a matrix and eigenvectors are special vectors.

Eigenvectors and eigenvalues

A matrix A *acts on* vectors $\mathbf{x}$ like a function does, with input $\mathbf{x}$ and output $A\mathbf{x}$. *Eigenvectors* are vectors for which $A\mathbf{x}$ is parallel to $\mathbf{x}$. In other words:

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

In this equation, $\mathbf{x}$ is an eigenvector of A and λ is an *eigenvalue* of A.

Eigenvalue 0

If the eigenvalue λ equals 0 then $A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. Vectors with eigenvalue 0 make up the nullspace of A; if A is singular, then $\lambda = 0$ is an eigenvalue of A.

Examples

Suppose P is the matrix of a projection onto a plane. For any $\mathbf{x}$ in the plane $P\mathbf{x} = \mathbf{x}$, so $\mathbf{x}$ is an eigenvector with eigenvalue 1. A vector $\mathbf{x}$ perpendicular to the plane has $P\mathbf{x} = \mathbf{0}$, so this is an eigenvector with eigenvalue $\lambda = 0$. The eigenvectors of P span the whole space (but this is not true for every matrix).

The matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has an eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue 1 and another eigenvector $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ with eigenvalue -1. These eigenvectors span the space. They are perpendicular because $B = B^T$ (as we will prove).

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

An n by n matrix will have n eigenvalues, and their sum will be the sum of the diagonal entries of the matrix: $a_{11} + a_{22} + \cdots + a_{nn}$. This sum is the *trace* of the matrix. For a two by two matrix, if we know one eigenvalue we can use this fact to find the second.

Can we solve $A\mathbf{x} = \lambda\mathbf{x}$ for the eigenvalues and eigenvectors of A? Both λ and $\mathbf{x}$ are unknown; we need to be clever to solve this problem:

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ (A - \lambda I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

In order for λ to be an eigenvector, $A - \lambda I$ must be singular. In other words, $\det(A - \lambda I) = 0$. We can solve this *characteristic equation* for λ to get n solutions.

If we're lucky, the solutions are distinct. If not, we have one or more *repeated eigenvalues*.

Once we've found an eigenvalue λ, we can use elimination to find the nullspace of $A - \lambda I$. The vectors in that nullspace are eigenvectors of A with eigenvalue λ.

Calculating eigenvalues and eigenvectors

Let $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Then:

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)^2 - 1 \\ &= \lambda^2 - 6\lambda + 8. \end{aligned}$$

Note that the coefficient 6 is the trace (sum of diagonal entries) and 8 is the determinant of A. In general, the eigenvalues of a two by two matrix are the solutions to:

$$\lambda^2 - \text{trace}(A) \cdot \lambda + \det A = 0.$$

Just as the trace is the sum of the eigenvalues of a matrix, the product of the eigenvalues of any matrix equals its determinant.

For $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, the eigenvalues are $\lambda_1 = 4$ and $\lambda_2 = 2$. We find the eigenvector $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for $\lambda_1 = 4$ in the nullspace of $A - \lambda_1 I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$.

x_2 will be in the nullspace of $A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. The nullspace is an entire line; x_2 could be any vector on that line. A natural choice is $x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Note that these eigenvectors are the same as those of $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Adding $3I$ to the matrix $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ added 3 to each of its eigenvalues and did not change its eigenvectors, because $A\mathbf{x} = (B + 3I)\mathbf{x} = \lambda\mathbf{x} + 3\mathbf{x} = (\lambda + 3)\mathbf{x}$.

A caution

Similarly, if $A\mathbf{x} = \lambda\mathbf{x}$ and $B\mathbf{x} = \alpha\mathbf{x}$, $(A + B)\mathbf{x} = (\lambda + \alpha)\mathbf{x}$. It would be nice if the eigenvalues of a matrix sum were always the sums of the eigenvalues, but this is only true if A and B have the same eigenvectors. The eigenvalues of the product AB aren't usually equal to the products $\lambda(A)\lambda(b)$, either.

Complex eigenvalues

The matrix $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotates every vector in the plane by 90°. It has trace $0 = \lambda_1 + \lambda_2$ and determinant $1 = \lambda_1 \cdot \lambda_2$. Its only real eigenvector is the zero vector; any other vector's direction changes when it is multiplied by Q. How will this affect our eigenvalue calculation?

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} \\ &= \lambda^2 + 1. \end{aligned}$$

$\det(A - \lambda I) = 0$ has solutions $\lambda_1 = i$ and $\lambda_2 = -i$. If a matrix has a complex eigenvalue $a + bi$ then the *complex conjugate* $a - bi$ is also an eigenvalue of that matrix.

Symmetric matrices have real eigenvalues. For *antisymmetric* matrices like Q, for which $A^T = -A$, all eigenvalues are imaginary ($\lambda = bi$).

Triangular matrices and repeated eigenvalues

For triangular matrices such as $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$, the eigenvalues are exactly the entries on the diagonal. In this case, the eigenvalues are 3 and 3:

$$\begin{aligned} \det(A - \lambda \det I) &= \begin{vmatrix} 3-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} \\ &= (3-\lambda)(3-\lambda) \quad \Big(= (a_{11} - \lambda)(a_{22} - \lambda) \Big) \\ &= 0, \end{aligned}$$

so $\lambda_1 = 3$ and $\lambda_2 = 3$. To find the eigenvectors, solve:

$$(A - \lambda I)\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

to get $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. There is no independent eigenvector $\mathbf{x}_2$.

Diagonalization and powers of A

We know how to find eigenvalues and eigenvectors. In this lecture we learn to *diagonalize* any matrix that has n independent eigenvectors and see how diagonalization simplifies calculations. The lecture concludes by using eigenvalues and eigenvectors to solve *difference equations*.

Diagonalizing a matrix $S^{-1}AS = \Lambda$

If A has n linearly independent eigenvectors, we can put those vectors in the columns of a (square, invertible) matrix S. Then

$$\begin{aligned} AS &= A\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \cdots & \lambda_n\mathbf{x}_n \end{bmatrix} \\ &= S\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = S\Lambda. \end{aligned}$$

Note that Λ is a diagonal matrix whose non-zero entries are the eigenvalues of A. Because the columns of S are independent, S^{-1} exists and we can multiply both sides of $AS = S\Lambda$ by S^{-1}:

$$S^{-1}AS = \Lambda.$$

Equivalently, $A = S\Lambda S^{-1}$.

Powers of A

What are the eigenvalues and eigenvectors of A^2?

$$\begin{aligned} \text{If } A\mathbf{x} &= \lambda\mathbf{x}, \\ \text{then } A^2\mathbf{x} &= \lambda A\mathbf{x} = \lambda^2\mathbf{x}. \end{aligned}$$

The eigenvalues of A^2 are the squares of the eigenvalues of A. The eigenvectors of A^2 are the same as the eigenvectors of A. If we write $A = S\Lambda S^{-1}$ then:

$$A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}.$$

Similarly, $A^k = S\Lambda^k S^{-1}$ tells us that raising the eigenvalues of A to the kth power gives us the eigenvalues of A^k, and that the eigenvectors of A^k are the same as those of A.

Theorem: If A has n independent eigenvectors with eigenvalues λ_i, then $A^k \to 0$ as $k \to \infty$ if and only if all $|\lambda_i| < 1$.

A is guaranteed to have n independent eigenvectors (and be *diagonalizable*) if all its eigenvalues are different. Most matrices do have distinct eigenvalues.

Repeated eigenvalues

If A has repeated eigenvalues, it may or may not have n independent eigenvectors. For example, the eigenvalues of the identity matrix are all 1, but that matrix still has n independent eigenvectors.

If A is the triangular matrix $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ its eigenvalues are 2 and 2. Its eigenvectors are in the nullspace of $A - \lambda I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ which is spanned by $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. This particular A does not have two independent eigenvectors.

Difference equations $\mathbf{u}_{k+1} = A\mathbf{u}_k$

Start with a given vector $\mathbf{u}_0$. We can create a sequence of vectors in which each new vector is A times the previous vector: $\mathbf{u}_{k+1} = A\mathbf{u}_k$. $\mathbf{u}_{k+1} = A\mathbf{u}_k$ is a *first order difference equation*, and $\mathbf{u}_k = A^k\mathbf{u}_0$ is a solution to this system.

We get a more satisfying solution if we write $\mathbf{u}_0$ as a combination of eigenvectors of A:

$$\mathbf{u}_0 = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n = S\mathbf{c}.$$

Then:

$$A\mathbf{u}_0 = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_n\lambda_n\mathbf{x}_n$$

and:

$$\mathbf{u}_k = A^k\mathbf{u}_0 = c_1\lambda_1^k\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n = \Lambda^k S\mathbf{c}.$$

Fibonacci sequence

The Fibonacci sequence is $0, 1, 1, 2, 3, 5, 8, 13, \ldots$. In general, $F_{k+2} = F_{k+1} + F_k$. If we could understand this in terms of matrices, the eigenvalues of the matrices would tell us how fast the numbers in the sequence are increasing.

$\mathbf{u}_{k+1} = A\mathbf{u}_k$ was a first order system. $F_{k+2} = F_{k+1} + F_k$ is a second order scalar equation, but we can convert it to first order linear system by using a clever trick. If $\mathbf{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$, then:

$$\begin{aligned} F_{k+2} &= F_{k+1} + F_k \quad (1) \\ F_{k+1} &= F_{k+1}. \quad (2) \end{aligned}$$

is equivalent to the first order system $\mathbf{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\mathbf{u}_k$.

What are the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$? Because A is symmetric, its eigenvalues will be real and its eigenvectors will be orthogonal.

Because A is a two by two matrix we know its eigenvalues sum to 1 (the trace) and their product is -1 (the determinant).

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 1$$

Setting this to zero we find $\lambda = \frac{1 \pm \sqrt{1+4}}{2}$; i.e. $\lambda_1 = \frac{1}{2}(1+\sqrt{5}) \approx 1.618$ and $\lambda_2 = \frac{1}{2}(1-\sqrt{5}) \approx -.618$. The growth rate of the F_k is controlled by λ_1, the only eigenvalue with absolute value greater than 1. This tells us that for large k, $F_k \approx c_1 \left(\frac{1+\sqrt{5}}{2}\right)^k$ for some constant c_1. (Remember $\mathbf{u}_k = A^k \mathbf{u}_0 = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2$, and here λ_2^k goes to zero since $|\lambda_2| < 1$.)

To find the eigenvectors of A note that:

$$(A - \lambda I)\mathbf{x} = \begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} \mathbf{x}$$

equals $\mathbf{0}$ when $\mathbf{x} = \begin{bmatrix} \lambda \\ 1 \end{bmatrix}$, so $\mathbf{x}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$.

Finally, $\mathbf{u}_0 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2$ tells us that $c_1 = -c_2 = \frac{1}{\sqrt{5}}$. Because $\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \mathbf{u}_k = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2$, we get:

$$F_k = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^k - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^k.$$

Using eigenvalues and eigenvectors, we have found a *closed form expression* for the Fibonacci numbers.

Summary: When a sequence evolves over time according to the rules of a first order system, the eigenvalues of the matrix of that system determine the long term behavior of the series. To get an exact formula for the series we find the eigenvectors of the matrix and then solve for the coefficients $c_1, c_2, ...$

Differential equations and e^{At}

The system of equations below describes how the values of variables u_1 and u_2 affect each other over time:

$$\frac{du_1}{dt} = -u_1 + 2u_2$$
$$\frac{du_2}{dt} = u_1 - 2u_2.$$

Just as we applied linear algebra to solve a difference equation, we can use it to solve this differential equation. For example, the initial condition $u_1 = 1$, $u_2 = 0$ can be written $\mathbf{u}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Differential equations $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$

By looking at the equations above, we might guess that over time u_1 will decrease. We can get the same sort of information more safely by looking at the eigenvalues of the matrix $A = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}$ of our system $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$. Because A is singular and its trace is -3 we know that its eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = -3$. The solution will turn out to include e^{-3t} and e^{0t}. As t increases, e^{-3t} vanishes and $e^{0t} = 1$ remains constant. Eigenvalues equal to zero have eigenvectors that are *steady state* solutions.

$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector for which $A\mathbf{x}_1 = 0\mathbf{x}_1$. To find an eigenvector corresponding to $\lambda_2 = -3$ we solve $(A - \lambda_2 I)\mathbf{x}_2 = \mathbf{0}$:

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \mathbf{x}_2 = 0 \quad \text{so} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and we can check that $A\mathbf{x}_2 = -3\mathbf{x}_2$. The general solution to this system of differential equations will be:

$$\mathbf{u}(t) = c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2.$$

Is $e^{\lambda_1 t}\mathbf{x}_1$ really a solution to $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$? To find out, plug in $\mathbf{u} = e^{\lambda_1 t}\mathbf{x}_1$:

$$\frac{d\mathbf{u}}{dt} = \lambda_1 e^{\lambda_1 t}\mathbf{x}_1,$$

which agrees with:

$$A\mathbf{u} = e^{\lambda_1 t}A\mathbf{x}_1 = \lambda_1 e^{\lambda_1 t}\mathbf{x}_1.$$

The two "pure" terms $e^{\lambda_1 t}\mathbf{x}_1$ and $e^{\lambda_2 t}\mathbf{x}_2$ are analogous to the terms $\lambda_i^k \mathbf{x}_i$ we saw in the solution $c_1\lambda_1^k\mathbf{x}_1 + c_2\lambda_2^k\mathbf{x}_2 + \cdots + c_n\lambda_n^k\mathbf{x}_n$ to the difference equation $\mathbf{u}_{k+1} = A\mathbf{u}_k$.

Plugging in the values of the eigenvectors, we get:

$$\mathbf{u}(t) = c_1 e^{\lambda_1 t}\mathbf{x}_1 + c_2 e^{\lambda_2 t}\mathbf{x}_2 = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We know $\mathbf{u}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so at $t = 0$:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$c_1 = c_2 = 1/3$ and $\mathbf{u}(t) = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3}e^{-3t}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

This tells us that the system starts with $u_1 = 1$ and $u_2 = 0$ but that as t approaches infinity, u_1 decays to $2/3$ and u_2 increases to $1/3$. This might describe stuff moving from u_1 to u_2.

The steady state of this system is $\mathbf{u}(\infty) = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$.

Stability

Not all systems have a steady state. The eigenvalues of A will tell us what sort of solutions to expect:

1. Stability: $\mathbf{u}(t) \to 0$ when $\mathrm{Re}(\lambda) < 0$.
2. Steady state: One eigenvalue is 0 and all other eigenvalues have negative real part.
3. Blow up: if $\mathrm{Re}(\lambda) > 0$ for any eigenvalue λ.

If a two by two matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has two eigenvalues with negative real part, its trace $a + d$ is negative. The converse is not true: $\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$ has negative trace but one of its eigenvalues is 1 and e^{1t} blows up. If A has a positive determinant and negative trace then the corresponding solutions must be stable.

Applying S

The final step of our solution to the system $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ was to solve:

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In matrix form:

$$\begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

or $S\mathbf{c} = \mathbf{u}(0)$, where S is the eigenvector matrix. The components of $\mathbf{c}$ determine the contribution from each pure exponential solution, based on the initial conditions of the system.

In the equation $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$, the matrix A couples the pure solutions. We set $\mathbf{u} = S\mathbf{v}$, where S is the matrix of eigenvectors of A, to get:

$$S\frac{d\mathbf{v}}{dt} = AS\mathbf{v}$$

or:

$$\frac{d\mathbf{v}}{dt} = S^{-1}AS\mathbf{v} = \Lambda\mathbf{v}.$$

This diagonalizes the system: $\frac{dv_i}{dt} = \lambda_i v_i$. The general solution is then:

$$\begin{aligned} \mathbf{v}(t) &= e^{\Lambda t}\mathbf{v}(0), \qquad \text{and} \\ \mathbf{u}(t) &= Se^{\Lambda t}S^{-1}\mathbf{v}(0) = e^{At}\mathbf{u}(0). \end{aligned}$$

Matrix exponential e^{At}

What does e^{At} mean if A is a matrix? We know that for a real number x,

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots.$$

We can use the same formula to define e^{At}:

$$e^{At} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots.$$

Similarly, if the eigenvalues of At are small, we can use the geometric series $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ to estimate $(I - At)^{-1} = I + At + (At)^2 + (At)^3 + \cdots$.

We've said that $e^{At} = Se^{\Lambda t}S^{-1}$. If A has n independent eigenvectors we can prove this from the definition of e^{At} by using the formula $A = S\Lambda S^{-1}$:

$$\begin{aligned} e^{At} &= I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots \\ &= SS^{-1} + S\Lambda S^{-1}t + \frac{S\Lambda^2 S^{-1}}{2}t^2 + \frac{S\Lambda^3 S^{-1}}{6}t^3 + \cdots \\ &= Se^{\Lambda t}S^{-1}. \end{aligned}$$

It's impractical to add up infinitely many matrices. Fortunately, there is an easier way to compute $e^{\Lambda t}$. Remember that:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

When we plug this in to our formula for e^{At} we find that:

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & e^{\lambda_n t} \end{bmatrix}.$$

This is another way to see the relationship between the stability of $\mathbf{u}(t) = Se^{\Lambda t}S^{-1}\mathbf{v}(0)$ and the eigenvalues of A.

Second order

We can change the second order equation $y'' + by' + ky = 0$ into a two by two first order system using a method similar to the one we used to find a formula for the Fibonacci numbers. If $u = \begin{bmatrix} y' \\ y \end{bmatrix}$, then

$$u' = \begin{bmatrix} y'' \\ y' \end{bmatrix} = \begin{bmatrix} -b & -k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y' \\ y \end{bmatrix}.$$

We could use the methods we just learned to solve this system, and that would give us a solution to the second order scalar equation we started with.

If we start with a kth order equation we get a k by k matrix with coefficients of the equation in the first row and 1's on a diagonal below that; the rest of the entries are 0.

Markov matrices; Fourier series

In this lecture we look at Markov matrices and Fourier series – two applications of eigenvalues and projections.

Eigenvalues of A^T

The eigenvalues of A and the eigenvalues of A^T are the same:

$$(A - \lambda I)^T = A^T - \lambda I,$$

so property 10 of determinants tells us that $\det(A - \lambda I) = \det(A^T - \lambda I)$. If λ is an eigenvalue of A then $\det(A^T - \lambda I) = 0$ and λ is also an eigenvalue of A^T.

Markov matrices

A matrix like:

$$A = \begin{bmatrix} .1 & .01 & .3 \\ .2 & .99 & .3 \\ .7 & 0 & .4 \end{bmatrix}$$

in which all entries are non-negative and each column adds to 1 is called a *Markov matrix*. These requirements come from Markov matrices' use in probability. Squaring or raising a Markov matrix to a power gives us another Markov matrix.

When dealing with systems of differential equations, eigenvectors with the eigenvalue 0 represented steady states. Here we're dealing with powers of matrices and get a steady state when $\lambda = 1$ is an eigenvalue.

The constraint that the columns add to 1 guarantees that 1 is an eigenvalue. All other eigenvalues will be less than 1. Remember that (if A has n independent eigenvectors) the solution to $\mathbf{u}_k = A^k \mathbf{u}_0$ is $\mathbf{u}_k = c_1 \lambda_1^k \mathbf{x}_1 + c_2 \lambda_2^k \mathbf{x}_2 + \cdots + c_n \lambda_n^k \mathbf{x}_n$. If $\lambda_1 = 1$ and all others eigenvalues are less than one the system approaches the steady state $c_1 \mathbf{x}_1$. This is the $\mathbf{x}_1$ component of $\mathbf{u}_0$.

Why does the fact that the columns sum to 1 guarantee that 1 is an eigenvalue? If 1 is an eigenvalue of A, then:

$$A - 1I = \begin{bmatrix} -.9 & .01 & .3 \\ .2 & -.01 & .3 \\ .7 & 0 & -.6 \end{bmatrix}$$

should be singular. Since we've subtracted 1 from each diagonal entry, the sum of the entries in each column of $A - I$ is zero. But then the sum of the rows of $A - I$ must be the zero row, and so $A - I$ is singular. The eigenvector $\mathbf{x}_1$ is in the nullspace of $A - I$ and has eigenvalue 1. It's not very hard to find $\mathbf{x}_1 = \begin{bmatrix} .6 \\ 33 \\ .7 \end{bmatrix}$.

We're studying the equation $\mathbf{u}_{k+1} = A\mathbf{u}_k$ where A is a Markov matrix. For example u_1 might be the population of (number of people in) Massachusetts and u_2 might be the population of California. A might describe what fraction of the population moves from state to state, or the probability of a single person moving. We can't have negative numbers of people, so the entries of A will always be positive. We want to account for all the people in our model, so the columns of A add to $1 = 100\%$.

For example:

$$\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_{t=k+1} = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} \begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_{t=k}$$

assumes that there's a 90% chance that a person in California will stay in California and only a 10% chance that she or he will move, while there's a 20% percent chance that a Massachusetts resident will move to California. If our initial conditions are $\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix}$, then after one move $\mathbf{u}_1 = A\mathbf{u}_0$ is:

$$\begin{bmatrix} u_{\text{Cal}} \\ u_{\text{Mass}} \end{bmatrix}_1 = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = \begin{bmatrix} 200 \\ 800 \end{bmatrix}.$$

For the next few values of k, the Massachusetts population will decrease and the California population will increase while the total population remains constant at 1000.

To understand the long term behavior of this system we'll need the eigenvectors and eigenvalues of $\begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix}$. We know that one eigenvalue is $\lambda_1 = 1$. Because the trace $.9 + .8 = 1.7$ is the sum of the eigenvalues, we see that $\lambda_2 = .7$.

Next we calculate the eigenvectors:

$$A - \lambda_1 I = \begin{bmatrix} -.1 & .2 \\ .1 & -.2 \end{bmatrix} \mathbf{x}_1 = \mathbf{0},$$

so we choose $\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The eigenvalue 1 corresponds to the steady state solution, and $\lambda_2 = .7 < 1$, so the system approaches a limit in which $2/3$ of 1000 people live in California and $1/3$ of 1000 people are in Massachusetts. This will be the limit from any starting vector $\mathbf{u}_0$.

To know how the population is distributed after a finite number of steps we look for an eigenvector corresponding to $\lambda_2 = .7$:

$$A - \lambda_2 I = \begin{bmatrix} .2 & .2 \\ .1 & .1 \end{bmatrix} \mathbf{x}_1 = \mathbf{0},$$

so let $\mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

From what we learned about difference equations we know that:

$$\mathbf{u}_k = c_1 1^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 (.7)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

When $k = 0$ we have:

$$\mathbf{u}_0 = \begin{bmatrix} 0 \\ 1000 \end{bmatrix} = c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

so $c_1 = \frac{1000}{3}$ and $c_2 = \frac{2000}{3}$.

In some applications Markov matrices are defined differently – their rows add to 1 rather than their columns. In this case, the calculations are the transpose of everything we've done here.

Fourier series and projections

Expansion with an orthonormal basis

If we have an orthonormal basis $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n$ then we can write any vector $\mathbf{v}$ as $\mathbf{v} = x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + \cdots + x_n\mathbf{q}_n$, where:

$$\mathbf{q}_i^T\mathbf{v} = x_1\mathbf{q}_i^T\mathbf{q}_1 + x_2\mathbf{q}_i^T\mathbf{q}_2 + \cdots + x_n\mathbf{q}_i^T\mathbf{q}_n = x_i.$$

Since $\mathbf{q}_i^T\mathbf{q}_j = 0$ unless $i = j$, this equation gives $x_i = \mathbf{q}_i^T\mathbf{v}$.

In terms of matrices, $\begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{v}$, or $Q\mathbf{x} = \mathbf{v}$. So $\mathbf{x} = Q^{-1}\mathbf{v}$. Because the q_i form an orthonormal basis, $Q^{-1} = Q^T$ and $\mathbf{x} = Q^T\mathbf{v}$. This is another way to see that $x_i = \mathbf{q}_i^T\mathbf{v}$.

Fourier series

The key idea above was that the basis of vectors $\mathbf{q}_i$ was orthonormal. Fourier series are built on this idea. We can describe a function $f(x)$ in terms of trigonometric functions:

$$f(x) = a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + b_2\sin 2x + \cdots.$$

This *Fourier series* is an infinite sum and the previous example was finite, but the two are related by the fact that the cosines and sines in the Fourier series are orthogonal.

We're now working in an infinite dimensional vector space. The vectors in this space are functions and the (orthogonal) basis vectors are 1, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, ...

What does "orthogonal" mean in this context? How do we compute a dot product or *inner product* in this vector space? For vectors in $\mathbb{R}^n$ the inner product is $\mathbf{v}^T\mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n$. Functions are described by a continuum of values $f(x)$ rather than by a discrete collection of components v_i. The best parallel to the vector dot product is:

$$f^T g = \int_0^{2\pi} f(x)g(x)\,dx.$$

We integrate from 0 to 2π because Fourier series are periodic:

$$f(x) = f(x + 2\pi).$$

The inner product of two basis vectors is zero, as desired. For example,

$$\int_0^{2\pi} \sin x \cos x\,dx = \frac{1}{2}(\sin x)^2\Big|_0^{2\pi} = 0.$$

How do we find a_0, a_1, etc. to find the coordinates or *Fourier coefficients* of a function in this space? The constant term a_0 is the average value of the function. Because we're working with an orthonormal basis, we can use the inner product to find the coefficients a_i.

$$\begin{aligned}\int_0^{2\pi} f(x)\cos x\,dx &= \int_0^{2\pi} (a_0 + a_1\cos x + b_1\sin x + a_2\cos 2x + \cdots)\cos x\,dx \\ &= 0 + \int_0^{2\pi} a_1\cos^2 x\,dx + 0 + 0 + \cdots \\ &= a_1\pi.\end{aligned}$$

We conclude that $a_1 = \dfrac{1}{\pi}\displaystyle\int_0^{2\pi} f(x)\cos x\,dx$. We can use the same technique to find any of the values a_i.

Symmetric matrices and positive definiteness

Symmetric matrices are good – their eigenvalues are real and each has a complete set of orthonormal eigenvectors. Positive definite matrices are even better.

Symmetric matrices

A *symmetric matrix* is one for which $A = A^T$. If a matrix has some special property (e.g. it's a Markov matrix), its eigenvalues and eigenvectors are likely to have special properties as well. For a symmetric matrix with real number entries, the eigenvalues are real numbers and it's possible to choose a complete set of eigenvectors that are perpendicular (or even orthonormal).

If A has n independent eigenvectors we can write $A = S\Lambda S^{-1}$. If A is symmetric we can write $A = Q\Lambda Q^{-1} = Q\Lambda Q^T$, where Q is an orthogonal matrix. Mathematicians call this the *spectral theorem* and think of the eigenvalues as the "spectrum" of the matrix. In mechanics it's called the *principal axis theorem*.

In addition, any matrix of the form $Q\Lambda Q^T$ will be symmetric.

Real eigenvalues

Why are the eigenvalues of a symmetric matrix real? Suppose A is symmetric and $A\mathbf{x} = \lambda\mathbf{x}$. Then we can conjugate to get $\overline{A}\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. If the entries of A are real, this becomes $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$. (This proves that complex eigenvalues of real valued matrices come in conjugate pairs.)

Now transpose to get $\overline{\mathbf{x}}^T A^T = \overline{\mathbf{x}}^T\overline{\lambda}$. Because A is symmetric we now have $\overline{\mathbf{x}}^T A = \overline{\mathbf{x}}^T\overline{\lambda}$. Multiplying both sides of this equation on the right by $\mathbf{x}$ gives:

$$\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T\overline{\lambda}\mathbf{x}.$$

On the other hand, we can multiply $A\mathbf{x} = \lambda\mathbf{x}$ on the left by $\overline{\mathbf{x}}^T$ to get:

$$\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T\lambda\mathbf{x}.$$

Comparing the two equations we see that $\overline{\mathbf{x}}^T\overline{\lambda}\mathbf{x} = \overline{\mathbf{x}}^T\lambda\mathbf{x}$ and, unless $\overline{\mathbf{x}}^T\mathbf{x}$ is zero, we can conclude $\lambda = \overline{\lambda}$ is real.

How do we know $\overline{\mathbf{x}}^T\mathbf{x} \neq 0$?

$$\overline{\mathbf{x}}^T\mathbf{x} = \begin{bmatrix} \overline{x}_1 & \overline{x}_2 & \cdots & \overline{x}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2.$$

If $\mathbf{x} \neq \mathbf{0}$ then $\overline{\mathbf{x}}^T\mathbf{x} \neq 0$.

With complex vectors, as with complex numbers, multiplying by the conjugate is often helpful.

Symmetric matrices with real entries have $A = A^T$, real eigenvalues, and perpendicular eigenvectors. If A has complex entries, then it will have real eigenvalues and perpendicular eigenvectors if and only if $A = \overline{A}^T$. (The proof of this follows the same pattern.)

Projection onto eigenvectors

If $A = A^T$, we can write:

$$\begin{aligned} A &= Q\Lambda Q^T \\ &= \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \\ &= \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T \end{aligned}$$

The matrix $\mathbf{q}_k \mathbf{q}_k^T$ is the projection matrix onto $\mathbf{q}_k$, so every symmetric matrix is a combination of perpendicular projection matrices.

Information about eigenvalues

If we know that eigenvalues are real, we can ask whether they are positive or negative. (Remember that the signs of the eigenvalues are important in solving systems of differential equations.)

For very large matrices A, it's impractical to compute eigenvalues by solving $|A - \lambda I| = 0$. However, it's not hard to compute the pivots, and the signs of the pivots of a symmetric matrix are the same as the signs of the eigenvalues:

$$\text{number of positive pivots} = \text{number of positive eigenvalues.}$$

Because the eigenvalues of $A + bI$ are just b more than the eigenvalues of A, we can use this fact to find which eigenvalues of a symmetric matrix are greater or less than any real number b. This tells us a lot about the eigenvalues of A even if we can't compute them directly.

Positive definite matrices

A *positive definite matrix* is a symmetric matrix A for which all eigenvalues are positive. A good way to tell if a matrix is positive definite is to check that all its pivots are positive.

Let $A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}$. The pivots of this matrix are 5 and $(\det A)/5 = 11/5$. The matrix is symmetric and its pivots (and therefore eigenvalues) are positive, so A is a positive definite matrix. Its eigenvalues are the solutions to:

$$|A - \lambda I| = \lambda^2 - 8\lambda + 11 = 0,$$

i.e. $4 \pm \sqrt{5}$.

The determinant of a positive definite matrix is always positive but the determinant of $\begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}$ is also positive, and that matrix isn't positive definite. If all of the subdeterminants of A are positive (determinants of the k by k matrices in the upper left corner of A, where $1 \leq k \leq n$), then A is positive definite.

The subject of positive definite matrices brings together what we've learned about pivots, determinants and eigenvalues of square matrices. Soon we'll have a chance to bring together what we've learned in this course and apply it to non-square matrices.

Complex matrices; fast Fourier transform

Matrices with all real entries can have complex eigenvalues! So we can't avoid working with complex numbers. In this lecture we learn to work with complex vectors and matrices.

The most important complex matrix is the Fourier matrix F_n, which is used for Fourier transforms. Normally, multiplication by F_n would require n^2 multiplications. The fast Fourier transform (FFT) reduces this to roughly $n \log_2 n$ multiplications, a revolutionary improvement.

Complex vectors

Length

Given a vector $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$ with complex entries, how do we find its length? Our old definition:

$$\mathbf{z}^T\mathbf{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

is no good; this quantity isn't always positive. For example:

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

We don't want to define the length of $\begin{bmatrix} 1 \\ i \end{bmatrix}$ to be 0. The correct definition is: $|\mathbf{z}|^2 = \overline{\mathbf{z}}^T\mathbf{z} = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2$. Then we have:

$$\left(\text{length} \begin{bmatrix} 1 \\ i \end{bmatrix}\right)^2 = \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = 2.$$

To simplify our notation we write $|\mathbf{z}|^2 = \mathbf{z}^H\mathbf{z}$, where $\mathbf{z}^H = \overline{\mathbf{z}}^T$. The H comes from the name Hermite, and $\mathbf{z}^H\mathbf{z}$ is read "$\mathbf{z}$ Hermitian $\mathbf{z}$".

Inner product

Similarly, the inner or dot product of two complex vectors is not just $\mathbf{y}^T\mathbf{x}$. We must also take the complex conjugate of $\mathbf{y}$:

$$\mathbf{y}^H\mathbf{x} = \overline{\mathbf{y}}^T\mathbf{x} = \overline{y}_1x_1 + \overline{y}_2x_2 + \cdots + \overline{y}_nx_n.$$

Complex matrices

Hermitian matrices

Symmetric matrices are real valued matrices for which $A^T = A$. If A is complex, a nicer property is $\overline{A}^T = A$; such a matrix is called *Hermitian* and we abbreviate $\overline{A}^T$ as A^H. Note that the diagonal entries of a Hermitian matrix must be real. For example,

$$\overline{A}^T = A = \begin{bmatrix} 2 & 3+i \\ 3-i & 5 \end{bmatrix}.$$

Similar to symmetric matrices, Hermitian matrices have real eigenvalues and perpendicular eigenvectors.

Unitary matrices

What does it mean for complex vectors $\mathbf{q}_1, \mathbf{q}_2, ..., \mathbf{q}_n$ to be perpendicular (or orthonormal)? We must use our new definition of the inner product. For a collection of $\mathbf{q}_j$ in complex space to be orthonormal, we require:

$$\overline{\mathbf{q}}_j \mathbf{q}_k = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}$$

We can again define $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$, and then $Q^H Q = I$. Just as "Hermitian" is the complex equivalent of "symmetric", the term "*unitary*" is analogous to "orthogonal". A *unitary matrix* is a square matrix with perpendicular columns of unit length.

Discrete Fourier transform

A *Fourier series* is a way of writing a periodic function or *signal* as a sum of functions of different frequencies:

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \cdots .$$

When working with finite data sets, the *discrete Fourier transform* is the key to this decomposition.

In electrical engineering and computer science, the rows and columns of a matrix are numbered starting with 0, not 1 (and ending with $n-1$, not n). We'll follow this convention when discussing the Fourier matrix:

$$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w & w^2 & & w^{n-1} \\ 1 & w^2 & w^4 & & w^{2(n-1)} \\ \vdots & & & \ddots & \vdots \\ 1 & w^{n-1} & w^{2(n-1)} & \cdots & w^{(n-1)^2} \end{bmatrix}.$$

Notice that $F_n = F_n^T$ and $(F_n)_{jk} = w^{jk}$, where $j, k = 0, 1, ..., n-1$ and the complex number w is $w = e^{i \cdot 2\pi/n}$ (so $w^n = 1$). The columns of this matrix are orthogonal.

All the entries of F_n are on the unit circle in the complex plane, and raising each one to the nth power gives 1. We could write $w = \cos(2\pi/n) + i\sin(2\pi/n)$, but that would just make it harder to compute w^{jk}.

Because $w^4 = 1$ and $w = e^{2\pi i/4} = i$, our best example of a Fourier matrix is:

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}.$$

To find the Fourier transform of a vector with four components (four data points) we multiply by F_4.

It's easy to check that the columns of F_4 are orthogonal, as long as we remember to conjugate when computing the inner product. However, F_4 is not quite unitary because each column has length 2. We could divide each entry by 2 to get a matrix whose columns are orthonormal:

$$\frac{1}{4} F_4^H F_4 = I.$$

An example

The signal corresponding to a single impulse at time zero is (roughly) described by $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. To find the Fourier transform of this signal we compute:

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

A single impulse has all frequencies in equal amounts.

If we multiply by F_4 again we almost get back to $(1, 0, 0, 0)$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Because $\frac{1}{\sqrt{n}} F_n$ is unitary, multiplying by F_n and dividing by the scalar n inverts the transform.

Fast Fourier transform

Fourier matrices can be broken down into chunks with lots of zero entries; Fourier probably didn't notice this. Gauss did, but didn't realize how significant a discovery this was.

There's a nice relationship between F_n and F_{2n} related to the fact that $w_{2n}^2 = w_n$:

$$F_{2n} = \begin{bmatrix} I & D \\ I & -D \end{bmatrix} \begin{bmatrix} F_n & 0 \\ 0 & F_n \end{bmatrix} P,$$

where D is a diagonal matrix and P is a $2n$ by $2n$ permutation matrix:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ & & & \vdots & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ & & & \vdots & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

So, a $2n$ sized Fourier transform F times $\mathbf{x}$ which we might think would require $(2n)^2 = 4n^2$ operations can instead be performed using two size n Fourier transforms ($2n^2$ operations) plus two very simple matrix multiplications which require on the order of n multiplications. The matrix P picks out the even components x_0, x_2, x_4, ... of a vector first, and then the odd ones – this calculation can be done very quickly.

Thus we can do a Fourier transform of size 64 on a vector by separating the vector into its odd and even components, performing a size 32 Fourier transform on each half of its components, then recombining the two halves through a process which involves multiplication by the diagonal matrix D.

$$D = \begin{bmatrix} 1 & & & & \\ & w & & & \\ & & w^2 & & \\ & & & \ddots & \\ & & & & w^{n-1} \end{bmatrix}.$$

Of course we can break each of those copies of F_{32} down into two copies of F_{16}, and so on. In the end, instead of using n^2 operations to multiply by F_n we get the same result using about $\frac{1}{2} n \log_2 n$ operations.

A typical case is $n = 1024 = 2^{10}$. Simply multiplying by F_n requires over a million calculations. The fast Fourier transform can be completed with only $\frac{1}{2} n \log_2 n = 5 \cdot 1024$ calculations. This is 200 times faster!

This is only possible because Fourier matrices are special matrices with orthogonal columns. In the next lecture we'll return to dealing exclusively with real numbers and will learn about positive definite matrices, which are the matrices most often seen in applications.

Positive definite matrices and minima

Studying positive definite matrices brings the whole course together; we use pivots, determinants, eigenvalues and stability. The new quantity here is $\mathbf{x}^T A\mathbf{x}$; watch for it.

This lecture covers how to tell if a matrix is positive definite, what it means for it to be positive definite, and some geometry.

Positive definite matrices

Given a symmetric two by two matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, here are four ways to tell if it's positive definite:

1. Eigenvalue test: $\lambda_1 > 0$, $\lambda_2 > 0$.
2. Determinants test: $a > 0$, $ac - b^2 > 0$.
3. Pivot test: $a > 0$, $\dfrac{ac - b^2}{a} > 0$.
4. $\mathbf{x}^T A\mathbf{x}$ is positive except when $\mathbf{x} = \mathbf{0}$ (this is usually the definition of positive definiteness).

2 by 2

Using the determinants test, we know that $\begin{bmatrix} 2 & 6 \\ 6 & y \end{bmatrix}$ is positive definite when $2y - 36 > 0$ or when $y > 18$.

The matrix $\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$ is on the borderline of positive definiteness and is called a *positive semidefinite* matrix. It's a singular matrix with eigenvalues 0 and 20. Positive semidefinite matrices have eigenvalues greater than or equal to 0. For a singular matrix, the determinant is 0 and it only has one pivot.

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2x_1 + 6x_2 \\ 6x_1 + 18x_2 \end{bmatrix} \\ &= 2x_1^2 + 12x_1x_2 + 18x_2^2 \\ &= ax_1^2 + 2bx_1x_2 + cx_2^2. \end{aligned}$$

If this *quadratic form* is positive for every (real) x_1 and x_2 then the matrix is positive definite. In this positive semi-definite example, $2x_1^2 + 12x_1x_2 + 18x_2^2 = 2(x_1 + 3x_2)^2 = 0$ when $x_1 = 3$ and $x_2 = -1$.

Tests for minimum

If we apply the fourth test to the matrix $\begin{bmatrix} 2 & 6 \\ 6 & 7 \end{bmatrix}$ which is not positive definite, we get the quadratic form $f(x,y) = 2x^2 + 12xy + 7y^2$. The graph of this function has a saddle point at the origin; see Figure 1.

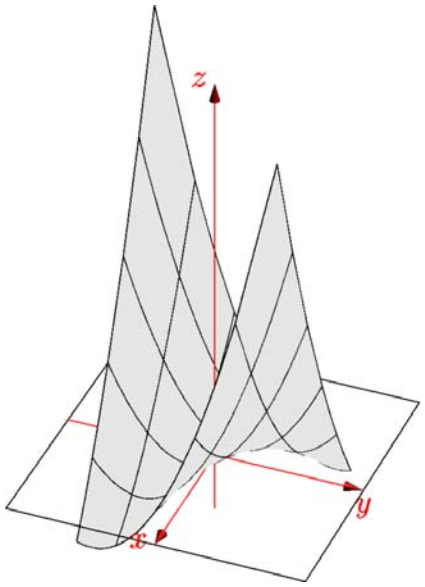

Figure 1: The graph of $f(x,y) = 2x^2 + 12xy + 7y^2$.

The matrix $\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix}$ is positive definite – its determinant is 4 and its trace is 22 so its eigenvalues are positive. The quadratic form associated with this matrix is $f(x,y) = 2x^2 + 12xy + 20y^2$, which is positive except when $x = y = 0$. The level curves $f(x,y) = k$ of this graph are ellipses; its graph appears in Figure 2. If $a > 0$ and $c > 0$, the quadratic form $ax^2 + 2bxy + cy^2$ is only negative when the value of $2bxy$ is negative and overwhelms the (positive) value of $ax^2 + cy^2$.

The first derivatives f_x and f_y of this function are zero, so its graph is tangent to the xy-plane at $(0,0,0)$; but this was also true of $2x^2 + 12xy + 7y^2$. As in single variable calculus, we need to look at the second derivatives of f to tell whether there is a minimum at the critical point.

We can prove that $2x^2 + 12xy + 20y^2$ is always positive by writing it as a sum of squares. We do this by completing the square:

$$2x^2 + 12xy + 20y^2 = 2(x + 3y)^2 + 2y^2.$$

Note that $2(x+3y)^2 = 2x^2 + 12xy + 18y^2$, and 18 was the "borderline" between passing and failing the tests for positive definiteness.

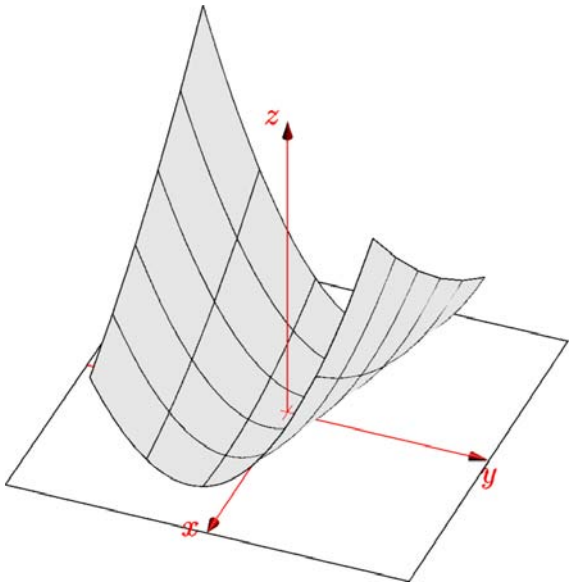

Figure 2: The graph of $f(x,y) = 2x^2 + 12xy + 20y^2$.

When we complete the square for $2x^2 + 12xy + 7y^2$ we get:

$$2x^2 + 12xy + 7y^2 = 2(x+3y)^2 - 11y^2$$

which may be negative; e.g. when $x = -3$ and $y = 1$.

The coefficients that appear when completing the square are exactly the entries that appear when performing elimination on the original matrix. The two pivots are multiplied by the squares, and the coefficient c in the term $(x - cy)^2$ is the multiple of the first row that's subtracted from the second row.

$$\begin{bmatrix} 2 & 6 \\ 6 & 20 \end{bmatrix} \xrightarrow{\text{subtract } \mathbf{3} \text{ times row 1}} \begin{bmatrix} \mathbf{2} & 6 \\ 0 & \mathbf{2} \end{bmatrix}.$$

We can see the terms that appear when completing the square in:

$$U = \begin{bmatrix} \mathbf{2} & 6 \\ 0 & \mathbf{2} \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 \\ \mathbf{3} & 1 \end{bmatrix}.$$

When we complete the square, the numbers multiplied by the squares are the pivots; if the pivots are all positive then the sum of squares will always be positive.

Hessian matrix

The matrix of second derivatives of $f(x,y)$ is:

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}.$$

This matrix is symmetric because $f_{xy} = f_{yx}$. Its determinant is positive when the matrix is positive definite, which matches the $f_{xx}f_{yy} > f_{xy}^2$ test for a minimum that we learned in calculus.

n by n

A function of several variables $f(x_1, x_2, ..., x_n)$ has a minimum when its matrix of second derivatives is positive definite, and identifying minima of functions is often important. The tests we've just learned for 2 by 2 matrices also apply to n by n matrices.

A 3 by 3 example:

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

Is this matrix positive definite? Our tests will say *yes*. What's the function $\mathbf{x}^T A\mathbf{x}$ associated with this matrix? Does that function have a minimum at $\mathbf{x} = \mathbf{0}$? What does the graph of its quadratic form look like?

Looking at determinants we see:

$$\det[\,2\,] = 2, \qquad \det\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 5, \qquad \det\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 4.$$

These are all positive, so A is positive definite.

The pivots of A are 2, $3/2$ and $4/3$ (all positive) because the products of the pivots equal the determinants.

The eigenvalues of A are positive and their product is 4. It's not difficult to check that they are $2 - \sqrt{2}$, 2 and $2 + \sqrt{2}$ (all positive).

Ellipsoids in $\mathbb{R}^n$

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3.$$

Because A is positive definite, we expect $f(\mathbf{x})$ to be positive except when $\mathbf{x} = \mathbf{0}$. Its graph is a sort of four dimensional bowl or *paraboloid*. If we wrote $f(\mathbf{x})$ as a sum of three squares, those squares would be multiplied by the (positive) pivots of A. Earlier, we said that a horizontal slice of our three dimensional bowl shape would be an ellipse. Here, a horizontal slice of the four dimensional bowl is an ellipsoid – a little bit like a rugby ball. For example, if we cut the graph at height 1 we get a surface whose equation is: $2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 = 1$.

Just as an ellipse has a major and minor axis, an ellipsoid has three axes. If we write $A = Q\Lambda Q^T$, as the principal axis theorem tells us we can, the eigenvectors of A tell us the directions of the principal axes of the ellipsoid. The eigenvalues tell us the lengths of those axes.

Similar matrices and Jordan form

We've nearly covered the entire heart of linear algebra – once we've finished singular value decompositions we'll have seen all the most central topics.

A^TA is positive definite

A matrix is *positive definite* if $\mathbf{x}^TA\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. This is a very important class of matrices; positive definite matrices appear in the form of A^TA when computing least squares solutions. In many situations, a rectangular matrix is multiplied by its transpose to get a square matrix.

Given a symmetric positive definite matrix A, is its inverse also symmetric and positive definite? Yes, because if the (positive) eigenvalues of A are $\lambda_1, \lambda_2, \cdots \lambda_d$ then the eigenvalues $1/\lambda_1, 1/\lambda_2, \cdots 1/\lambda_d$ of A^{-1} are also positive.

If A and B are positive definite, is $A + B$ positive definite? We don't know much about the eigenvalues of $A + B$, but we can use the property $\mathbf{x}^TA\mathbf{x} > 0$ and $\mathbf{x}^TB\mathbf{x} > 0$ to show that $\mathbf{x}^T(A + B)\mathbf{x} > 0$ for $\mathbf{x} \neq 0$ and so $A + B$ is also positive definite.

Now suppose A is a rectangular (m by n) matrix. A is almost certainly not symmetric, but A^TA is square and symmetric. Is A^TA positive definite? We'd rather not try to find the eigenvalues or the pivots of this matrix, so we ask when $\mathbf{x}^TA^TA\mathbf{x}$ is positive.

Simplifying $\mathbf{x}^TA^TA\mathbf{x}$ is just a matter of moving parentheses:

$$\mathbf{x}^T(A^TA)\mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = |A\mathbf{x}|^2 \geq 0.$$

The only remaining question is whether $A\mathbf{x} = \mathbf{0}$. If A has rank n (independent columns), then $\mathbf{x}^T(A^TA)\mathbf{x} = \mathbf{0}$ only if $\mathbf{x} = \mathbf{0}$ and A is positive definite.

Another nice feature of positive definite matrices is that you never have to do row exchanges when row reducing – there are never 0's or unsuitably small numbers in their pivot positions.

Similar matrices A and $B = M^{-1}AM$

Two square matrices A and B are *similar* if $B = M^{-1}AM$ for some matrix M. This allows us to put matrices into families in which all the matrices in a family are similar to each other. Then each family can be represented by a diagonal (or nearly diagonal) matrix.

Distinct eigenvalues

If A has a full set of eigenvectors we can create its eigenvector matrix S and write $S^{-1}AS = \Lambda$. So A is similar to Λ (choosing M to be this S).

If $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ then $\Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ and so A is similar to $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$. But A is also similar to:

$$\begin{matrix} M^{-1} \\ \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \end{matrix} \begin{matrix} A \\ \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \end{matrix} \begin{matrix} M \\ \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} \end{matrix} = \begin{matrix} B \\ \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 9 \\ 1 & 6 \end{bmatrix} \end{matrix}$$
$$= \begin{bmatrix} -2 & -15 \\ 1 & 6 \end{bmatrix}.$$

In addition, B is similar to Λ. All these similar matrices have the same eigenvalues, 3 and 1; we can check this by computing the trace and determinant of A and B.

Similar matrices have the same eigenvalues!

In fact, the matrices similar to A are all the 2 by 2 matrices with eigenvalues 3 and 1. Some other members of this family are $\begin{bmatrix} 3 & 7 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 7 \\ 0 & 3 \end{bmatrix}$. To prove that similar matrices have the same eigenvalues, suppose $A\mathbf{x} = \lambda\mathbf{x}$. We modify this equation to include $B = M^{-1}AM$:

$$\begin{aligned} AMM^{-1}\mathbf{x} &= \lambda\mathbf{x} \\ M^{-1}AMM^{-1}\mathbf{x} &= \lambda M^{-1}\mathbf{x} \\ BM^{-1}\mathbf{x} &= \lambda M^{-1}\mathbf{x}. \end{aligned}$$

The matrix B has the same λ as an eigenvalue. $M^{-1}\mathbf{x}$ is the eigenvector.

If two matrices are similar, they have the same eigenvalues and the same number of independent eigenvectors (but probably not the same eigenvectors).

When we diagonalize A, we're finding a diagonal matrix Λ that is similar to A. If two matrices have the same n distinct eigenvalues, they'll be similar to the same diagonal matrix.

Repeated eigenvalues

If two eigenvalues of A are the same, it may not be possible to diagonalize A. Suppose $\lambda_1 = \lambda_2 = 4$. One family of matrices with eigenvalues 4 and 4 contains only the matrix $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. The matrix $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ is not in this family.

There are two families of similar matrices with eigenvalues 4 and 4. The larger family includes $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$. Each of the members of this family has only one eigenvector.

The matrix $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ is the only member of the other family, because:

$$M^{-1}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}M = 4M^{-1}M = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

for any invertible matrix M.

Jordan form

Camille Jordan found a way to choose a "most diagonal" representative from each family of similar matrices; this representative is said to be in *Jordan normal form*. For example, both $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ and $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ are in Jordan form. This form used to be the climax of linear algebra, but not any more. Numerical applications rarely need it.

We can find more members of the family represented by $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ by choosing diagonal entries to get a trace of 4, then choosing off-diagonal entries to get a determinant of 16:

$$\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 5 & 1 \\ -1 & 3 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 17 & 4 \end{bmatrix}, \begin{bmatrix} a & b \\ (8a - a^2 - 16)/b & 8 - a \end{bmatrix}.$$

(None of these are diagonalizable, because if they were they would be similar to $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$. That matrix is only similar to itself.) What about this one?

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

Its eigenvalues are four zeros. Its rank is 2 so the dimension of its nullspace is $4 - 2 = 2$. It will have two independent eigenvectors and two "missing" eigenvectors. When we look instead at

$$\begin{bmatrix} 0 & 1 & 7 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

its rank and the dimension of its nullspace are still 2, but it's not as nice as A. B *is* similar to A, which is the Jordan normal form representative of this family. A has a 1 above the diagonal for every missing eigenvector and the rest of its entries are 0.

Now consider:

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

Again it has rank 2 and its nullspace has dimension 2. Its four eigenvalues are 0. Surprisingly, it is not similar to A. We can see this by breaking the matrices

into their *Jordan blocks*:

$$A = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right], \qquad C = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

A Jordan block J_i has a repeated eigenvalue λ_i on the diagonal, zeros below the diagonal and in the upper right hand corner, and ones above the diagonal:

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Two matrices may have the same eigenvalues and the same number of eigenvectors, but if their Jordan blocks are different sizes those matrices can not be similar.

Jordan's theorem says that every square matrix A is similar to a Jordan matrix J, with Jordan blocks on the diagonal:

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_d \end{bmatrix}.$$

In a Jordan matrix, the eigenvalues are on the diagonal and there may be ones above the diagonal; the rest of the entries are zero. The number of blocks is the number of eigenvectors – there is one eigenvector per block.

To summarize:

- If A has n distinct eigenvalues, it is diagonalizable and its Jordan matrix is the diagonal matrix $J = \Lambda$.
- If A has repeated eigenvalues and "missing" eigenvectors, then its Jordan matrix will have $n - d$ ones above the diagonal.

We have not learned to compute the Jordan matrix of a matrix which is missing eigenvectors, but we do know how to diagonalize a matrix which has n distinct eigenvalues.

Singular value decomposition

The *singular value decomposition* of a matrix is usually referred to as the *SVD*. This is the final and best factorization of a matrix:

$$A = U\Sigma V^T$$

where U is orthogonal, Σ is diagonal, and V is orthogonal.

In the decomoposition $A = U\Sigma V^T$, A can be *any* matrix. We know that if A is symmetric positive definite its eigenvectors are orthogonal and we can write $A = Q\Lambda Q^T$. This is a special case of a SVD, with $U = V = Q$. For more general A, the SVD requires two different matrices U and V.

We've also learned how to write $A = S\Lambda S^{-1}$, where S is the matrix of n distinct eigenvectors of A. However, S may not be orthogonal; the matrices U and V in the SVD will be.

How it works

We can think of A as a linear transformation taking a vector $\mathbf{v}_1$ in its row space to a vector $\mathbf{u}_1 = A\mathbf{v}_1$ in its column space. The SVD arises from finding an orthogonal basis for the row space that gets transformed into an orthogonal basis for the column space: $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$.

It's not hard to find an orthogonal basis for the row space – the Gram-Schmidt process gives us one right away. But in general, there's no reason to expect A to transform that basis to another orthogonal basis.

You may be wondering about the vectors in the nullspaces of A and A^T. These are no problem – zeros on the diagonal of Σ will take care of them.

Matrix language

The heart of the problem is to find an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_r$ for the row space of A for which

$$A\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \end{bmatrix} = \begin{bmatrix} \sigma_1\mathbf{u}_1 & \sigma_2\mathbf{u}_2 & \cdots & \sigma_r\mathbf{u}_r \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix},$$

with $\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_r$ an orthonormal basis for the column space of A. Once we add in the nullspaces, this equation will become $AV = U\Sigma$. (We can complete the orthonormal bases $\mathbf{v}_1, ...\mathbf{v}_r$ and $\mathbf{u}_1, ...\mathbf{u}_r$ to orthonormal bases for the entire space any way we want. Since $\mathbf{v}_{r+1}, ...\mathbf{v}_n$ will be in the nullspace of A, the diagonal entries $\sigma_{r+1}, ...\sigma_n$ will be 0.)

The columns of U and V are bases for the row and column spaces, respectively. Usually $U \neq V$, but if A is positive definite we can use the *same* basis for its row and column space!

Calculation

Suppose A is the invertible matrix $\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$. We want to find vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in the row space $\mathbb{R}^2$, $\mathbf{u}_1$ and $\mathbf{u}_2$ in the column space $\mathbb{R}^2$, and positive numbers σ_1 and σ_2 so that the $\mathbf{v}_i$ are orthonormal, the $\mathbf{u}_i$ are orthonormal, and the σ_i are the scaling factors for which $A\mathbf{v}_i = \sigma_i u_i$.

This is a big step toward finding orthonormal matrices V and U and a diagonal matrix Σ for which:

$$AV = U\Sigma.$$

Since V is orthogonal, we can multiply both sides by $V^{-1} = V^T$ to get:

$$A = U\Sigma V^T.$$

Rather than solving for U, V and Σ simultaneously, we multiply both sides by $A^T = V\Sigma^T U^T$ to get:

$$\begin{aligned} A^T A &= V\Sigma U^{-1} U\Sigma V^T \\ &= V\Sigma^2 V^T \\ &= V \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix} V^T. \end{aligned}$$

This is in the form $Q\Lambda Q^T$; we can now find V by diagonalizing the symmetric positive definite (or semidefinite) matrix $A^T A$. The columns of V are eigenvectors of $A^T A$ and the eigenvalues of $A^T A$ are the values σ_i^2. (We choose σ_i to be the positive square root of λ_i.)

To find U, we do the same thing with AA^T.

SVD example

We return to our matrix $A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}$. We start by computing

$$\begin{aligned} A^T A &= \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}. \end{aligned}$$

The eigenvectors of this matrix will give us the vectors $\mathbf{v}_i$, and the eigenvalues will gives us the numbers σ_i.

Two orthogonal eigenvectors of A^TA are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. To get an orthonormal basis, let $\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$. These have eigenvalues $\sigma_1^2 = 32$ and $\sigma_2^2 = 18$. We now have:

$$\begin{matrix} A \\ \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \end{matrix} = \begin{matrix} U \\ \begin{bmatrix} & \\ & \end{bmatrix} \end{matrix} \begin{matrix} \Sigma \\ \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix} \end{matrix} \begin{matrix} V^T \\ \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \end{matrix}.$$

We could solve this for U, but for practice we'll find U by finding orthonormal eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ for $AA^T = U\Sigma^2U^T$.

$$\begin{aligned} AA^T &= \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}. \end{aligned}$$

Luckily, AA^T happens to be diagonal. It's tempting to let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, as Professor Strang did in the lecture, but because $A\mathbf{v}_2 = \begin{bmatrix} 0 \\ -3\sqrt{2} \end{bmatrix}$ we instead have $\mathbf{u}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Note that this also gives us a chance to double check our calculation of σ_1 and σ_2.

Thus, the SVD of A is:

$$\begin{matrix} A \\ \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \end{matrix} = \begin{matrix} U \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{matrix} \begin{matrix} \Sigma \\ \begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix} \end{matrix} \begin{matrix} V^T \\ \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \end{matrix}.$$

Example with a nullspace

Now let $A = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$. This has a one dimensional nullspace and one dimensional row and column spaces.

The row space of A consists of the multiples of $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$. The column space of A is made up of multiples of $\begin{bmatrix} 4 \\ 8 \end{bmatrix}$. The nullspace and left nullspace are perpendicular to the row and column spaces, respectively.

Unit basis vectors of the row and column spaces are $\mathbf{v}_1 = \begin{bmatrix} .8 \\ .6 \end{bmatrix}$ and $\mathbf{u}_1 =$

$\begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$. To compute σ_1 we find the nonzero eigenvalue of A^TA.

$$A^TA = \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}$$
$$= \begin{bmatrix} 80 & 60 \\ 60 & 45 \end{bmatrix}.$$

Because this is a rank 1 matrix, one eigenvalue must be 0. The other must equal the trace, so $\sigma_1^2 = 125$. After finding unit vectors perpendicular to $\mathbf{u}_1$ and $\mathbf{v}_1$ (basis vectors for the left nullspace and nullspace, respectively) we see that the SVD of A is:

$$\underset{A}{\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix}} = \underset{U}{\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} \sqrt{125} & 0 \\ 0 & 0 \end{bmatrix}} \underset{V^T}{\begin{bmatrix} .8 & .6 \\ .6 & -.8 \end{bmatrix}}.$$

The singular value decomposition combines topics in linear algebra ranging from positive definite matrices to the four fundamental subspaces.

$\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_r$	is an orthonormal basis for the row space.
$\mathbf{u}_1, \mathbf{u}_2, ...\mathbf{u}_r$	is an orthonormal basis for the column space.
$\mathbf{v}_{r+1}, ...\mathbf{v}_n$	is an orthonormal basis for the nullspace.
$\mathbf{u}_{r+1}, ...\mathbf{u}_m$	is an orthonormal basis for the left nullspace.

These are the "right" bases to use, because $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$.

Linear transformations and their matrices

In older linear algebra courses, linear transformations were introduced before matrices. This geometric approach to linear algebra initially avoids the need for coordinates. But eventually there must be coordinates and matrices when the need for computation arises.

Without coordinates (no matrix)

Example 1: Projection

We can describe a projection as a *linear transformation* T which takes every vector in $\mathbb{R}^2$ into another vector in $\mathbb{R}^2$. In other words,

$$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

The rule for this *mapping* is that every vector $\mathbf{v}$ is projected onto a vector $T(\mathbf{v})$ on the line of the projection. Projection is a linear transformation.

Definition of *linear*

A transformation T is *linear* if:

$$T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$$

and

$$T(c\mathbf{v}) = cT(\mathbf{v})$$

for all vectors $\mathbf{v}$ and $\mathbf{w}$ and for all scalars c. Equivalently,

$$T(c\mathbf{v} + d\mathbf{w}) = cT(\mathbf{v}) + dT(\mathbf{w})$$

for all vectors $\mathbf{v}$ and $\mathbf{w}$ and scalars c and d. It's worth noticing that $T(\mathbf{0}) = \mathbf{0}$, because if not it couldn't be true that $T(c\mathbf{0}) = cT(\mathbf{0})$.

Non-example 1: Shift the whole plane

Consider the transformation $T(\mathbf{v}) = \mathbf{v} + \mathbf{v}_0$ that shifts every vector in the plane by adding some fixed vector $\mathbf{v}_0$ to it. This is *not* a linear transformation because $T(2\mathbf{v}) = 2\mathbf{v} + \mathbf{v}_0 \neq 2T(\mathbf{v})$.

Non-example 2: $T(\mathbf{v}) = ||\mathbf{v}||$

The transformation $T(\mathbf{v}) = ||\mathbf{v}||$ that takes any vector to its length is not a linear transformation because $T(c\mathbf{v}) \neq cT(\mathbf{v})$ if $c < 0$.

We're not going to study transformations that aren't linear. From here on, we'll only use T to stand for linear transformations.

Example 2: Rotation by 45°

This transformation $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ takes an input vector $\mathbf{v}$ and outputs the vector $T(\mathbf{v})$ that comes from rotating $\mathbf{v}$ counterclockwise by 45° about the origin. Note that we can describe this and see that it's linear without using any coordinates.

The big picture

One advantage of describing transformations geometrically is that it helps us to see the big picture, as opposed to focusing on the transformation's effect on a single point. We can quickly see how rotation by 45° will transform a picture of a house in the plane. If the transformation was described in terms of a matrix rather than as a rotation, it would be harder to guess what the house would be mapped to.

Frequently, the best way to understand a linear transformation is to find the matrix that lies behind the transformation. To do this, we have to choose a basis and bring in coordinates.

With coordinates (matrix!)

All of the linear transformations we've discussed above can be described in terms of matrices. In a sense, linear transformations are an abstract description of multiplication by a matrix, as in the following example.

Example 3: $T(\mathbf{v}) = A\mathbf{v}$

Given a matrix A, define $T(\mathbf{v}) = A\mathbf{v}$. This is a linear transformation:

$$A(\mathbf{v} + \mathbf{w}) = A(\mathbf{v}) + A(\mathbf{w})$$

and

$$A(c\mathbf{v}) = cA(\mathbf{v}).$$

Example 4

Suppose $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. How would we describe the transformation $T(\mathbf{v}) = A\mathbf{v}$ geometrically?

When we multiply A by a vector $\mathbf{v}$ in $\mathbb{R}^2$, the x component of the vector is unchanged and the sign of the y component of the vector is reversed. The transformation $\mathbf{v} \mapsto A\mathbf{v}$ reflects the xy-plane across the x axis.

Example 5

How could we find a linear transformation $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ that takes three dimensional space to two dimensional space? Choose any 2 by 3 matrix A and define $T(\mathbf{v}) = A\mathbf{v}$.

Describing $T(\mathbf{v})$

How much information do we need about T to to determine $T(\mathbf{v})$ for all $\mathbf{v}$? If we know how T transforms a single vector $\mathbf{v}_1$, we can use the fact that T is a linear transformation to calculate $T(c\mathbf{v}_1)$ for any scalar c. If we know $T(\mathbf{v}_1)$ and $T(\mathbf{v}_2)$ for two independent vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, we can predict how T will transform any vector $c\mathbf{v}_1 + d\mathbf{v}_2$ in the plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$. If we wish to know $T(\mathbf{v})$ for all vectors $\mathbf{v}$ in $\mathbb{R}^n$, we just need to know $T(\mathbf{v}_1), T(\mathbf{v}_2), ..., T(\mathbf{v}_n)$ for any basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ of the input space. This is because any $\mathbf{v}$ in the input space can be written as a linear combination of basis vectors, and we know that T is linear:

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

This is how we get from a (coordinate-free) linear transformation to a (coordinate based) matrix; the c_i are our coordinates. Once we've chosen a basis, every vector $\mathbf{v}$ in the space can be written as a combination of basis vectors in exactly one way. The coefficients of those vectors are the *coordinates* of $\mathbf{v}$ in that basis.

Coordinates come from a basis; changing the basis changes the coordinates of vectors in the space. We may not use the standard basis all the time – we sometimes want to use a basis of eigenvectors or some other basis.

The matrix of a linear transformation

Given a linear transformation T, how do we construct a matrix A that represents it?

First, we have to choose two bases, say $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ of $\mathbb{R}^n$ to give coordinates to the input vectors and $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m$ of $\mathbb{R}^m$ to give coordinates to the output vectors. We want to find a matrix A so that $T(\mathbf{v}) = A\mathbf{v}$, where $\mathbf{v}$ and $A\mathbf{v}$ get their coordinates from these bases.

The first column of A consists of the coefficients $a_{11}, a_{21}, ..., a_{1m}$ of $T(\mathbf{v}_1) = a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{1m}\mathbf{w}_m$. The entries of column i of the matrix A are determined by $T(\mathbf{v}_i) = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \cdots + a_{1i}\mathbf{w}_m$. Because we've guaranteed that $T(\mathbf{v}_i) = A\mathbf{v}_i$ for each basis vector $\mathbf{v}_i$ and because T is linear, we know that $T(\mathbf{v}) = A\mathbf{v}$ for all vectors $\mathbf{v}$ in the input space.

In the example of the projection matrix, $n = m = 2$. The transformation T projects every vector in the plane onto a line. In this example, it makes sense to use the same basis for the input and the output. To make our calculations as simple as possible, we'll choose $\mathbf{v}_1$ to be a unit vector on the line of projection and $\mathbf{v}_2$ to be a unit vector perpendicular to $\mathbf{v}_1$. Then

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{v}_1 + \mathbf{0}$$

and the matrix of the projection transformation is just $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

$$A\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ 0 \end{bmatrix}.$$

This is a nice matrix! If our chosen basis consists of eigenvectors then the matrix of the transformation will be the diagonal matrix Λ with eigenvalues on the diagonal.

To see how important the choice of basis is, let's use the standard basis for the linear transformation that projects the plane onto a line at a 45° angle. If we choose $\mathbf{v}_1 = \mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we get the projection matrix $P = \dfrac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$. We can check by graphing that this is the correct matrix, but calculating P directly is more difficult for this basis than it was with a basis of eigenvectors.

Example 6: $T = \dfrac{d}{dx}$

Let T be a transformation that takes the derivative:

$$T(c_1 + c_2x + c_3x^2) = c_2 + 2c_3x. \tag{1}$$

The input space is the three dimensional space of quadratic polynomials $c_1 + c_2x + c_3x^2$ with basis $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = x$ and $\mathbf{v}_3 = x^2$. The output space is a two dimensional subspace of the input space with basis $\mathbf{w}_1 = \mathbf{v}_1 = 1$ and $\mathbf{w}_2 = \mathbf{v}_2 = x$.

This is a linear transformation! So we can find $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ and write the transformation (1) as a matrix multiplication (2):

$$T\left(\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}\right) = A \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_2 \\ 2c_3 \end{bmatrix}. \tag{2}$$

Conclusion

For any linear transformation T we can find a matrix A so that $T(\mathbf{v}) = A\mathbf{v}$. If the transformation is invertible, the inverse transformation has the matrix A^{-1}. The product of two transformations $T_1 : \mathbf{v} \mapsto A_1\mathbf{v}$ and $T_2 : \mathbf{w} \mapsto A_2\mathbf{w}$ corresponds to the product A_2A_1 of their matrices. This is where matrix multiplication came from!

Change of basis; image compression

We've learned that computations can be made easier by an appropriate choice of basis. One application of this principle is to image compression. Lecture videos, music, and other data sources contain a lot of information; that information can be efficiently stored and transmitted only after we change the basis used to record it.

Compression of images

Suppose one frame of our lecture video is 512 by 512 pixels and that the video is recorded in black and white. The camera records a brightness level for each of the $(512)^2$ pixels; in this sense, each frame of video is a vector in a $(512)^2$ dimensional vector space.

The standard basis for this space has a vector for each pixel. Transmitting the values of all $(512)^2$ components of each frame using the standard basis would require far too much bandwidth, but if we change our basis according to the JPEG image compression standard we can transmit a fairly good copy the video very efficiently.

For example, if we're reporting light levels pixel by pixel, there's no efficient way to transmit the information "the entire frame is black". However, if one of our basis vectors corresponds to all pixels having the same light level (say 1), we can very efficiently transmit a recording of a blank blackboard.

Along with a vector of all 1's, we might choose a basis vector that alternates 1's and -1's, or one that's half 1's and half -1's corresponding to an image that's bright on the left and dark on the right. Our choice of basis will directly affect how much data we need to download to watch a video, and the best choice of basis for algebra lectures might differ from the best choice for action movies!

Fourier basis vectors

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \omega \\ \omega^2 \\ \omega^3 \\ \omega^4 \\ \omega^5 \\ \omega^6 \\ \omega^7 \end{bmatrix}, \begin{bmatrix} 1 \\ \omega^2 \\ \omega^4 \\ \omega^6 \\ \omega^8 \\ \omega^{10} \\ \omega^{12} \\ \omega^{14} \end{bmatrix}, \cdots, \begin{bmatrix} 1 \\ \omega^7 \\ \omega^{14} \\ \omega^{21} \\ \omega^{28} \\ \omega^{35} \\ \omega^{42} \\ \omega^{49} \end{bmatrix}$$

The best known basis is the Fourier basis, which is closely related to the Fourier matrices we studied earlier. The basis used by JPEG is made up of cosines – the real parts of ω^{jk}.

This method breaks the 512 by 512 rectangle of pixels into blocks that are 8 pixels on a side, each block containing 64 pixels total. The brightness information for those pixels is then compressed, possibly by eliminating all coefficients

below some threshold chosen so that we can hardly see the difference once they're gone.

$$\text{signal } \mathbf{x} \overset{\text{lossless}}{\longrightarrow} 64 \text{ coefficients } c \overset{\text{lossy compression}}{\longrightarrow} \hat{c} \text{ (many zeros)} \longrightarrow \hat{\mathbf{x}} = \sum \hat{c}_i \mathbf{v}_i$$

In video, not only should we consider compressing each frame, we can also consider compressing sequences of frames. There's very little difference between one frame and the next. If we do it right, we only need to encode and compress the differences between frames, not every frame in its entirety.

The Haar wavelet basis

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

The closest competitor to the JPEG encoding method uses a wavelet basis. (JPEG2000 improves on the Haar wavelets above.) In Haar's wavelet basis for $\mathbb{R}^8$, the non-zero entries are half 1's and half -1's (except for the vector of all 1's). However, half or even three quarters of a basis vector's entries may be 0. These vectors are chosen to be orthogonal and can be adjusted to be orthonormal.

Compression and matrices

Linear algebra is used to find the coefficients c_i in the change of basis from the standard basis (light levels for each pixel) to the Fourier or wavelet basis. For example, we might want to write:

$$\mathbf{x} = c_1 \mathbf{w}_1 + \cdots + c_8 \mathbf{w}_8.$$

But this is just a linear combination of the wavelet basis vectors. If W is the matrix whose columns are the wavelet vectors, then our task is simply to solve for $\mathbf{c}$:

$$\mathbf{x} = W \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_8 \end{bmatrix}.$$

So $\mathbf{c} = W^{-1}\mathbf{x}$.

Our calculations will be faster and easier if we don't have to spend a lot of time inverting a matrix (e.g. if $W^{-1} = W^T$) or multiplying by the inverse. So in the field of image compression, the criteria for a good basis are:

- Multiplication by the basis matrix and its inverse is fast (as in the FFT or in the wavelet basis).
- Good compression – the image can be approximated using only a few basis vectors. Most components c_i are small – safely set to zero.

Change of basis

Vectors

Let the columns of matrix W be the basis vectors of the new basis. Then if $\mathbf{x}$ is a vector in the old basis, we can convert it to a vector $\mathbf{c}$ in the new basis using the relation:

$$\mathbf{x} = W\mathbf{c}.$$

Transformation matrices

Suppose we have a linear transformation T. If T has the matrix A when working with the basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_8$ and T has the matrix B when working with the basis $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_8$, it turns out that A and B must be similar matrices. In other words, $B = M^{-1}AM$ for some change of basis matrix M.

Reminder: If we have a basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_8$ and we know $T(\mathbf{v}_i)$ for each i, then we can use the fact that T is a linear transformation to find $T(\mathbf{v}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_8T(\mathbf{v}_8)$ for any vector $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_8\mathbf{v}_8$ in the space. The entries of column i of the matrix A are the coefficients of the output vector $T(\mathbf{v}_i)$.

If our basis consists of eigenvectors of our transformation, i.e. if $T(\mathbf{v}_i) = \lambda_i\mathbf{v}_i$, then $A = \Lambda$, the (diagonal) matrix of eigenvalues. It would be wonderful to use a basis of eigenvectors for image processing, but finding such a basis requires far more computation than simply using a Fourier or wavelet basis.

Summary

When we change bases, the coefficients of our vectors change according to the rule $\mathbf{x} = W\mathbf{c}$. Matrix entries change according to a rule $B = M^{-1}AM$.

Left and right inverses; pseudoinverse

Although pseudoinverses will not appear on the exam, this lecture will help us to prepare.

Two sided inverse

A 2-*sided inverse* of a matrix A is a matrix A^{-1} for which $AA^{-1} = I = A^{-1}A$. This is what we've called the *inverse* of A. Here $r = n = m$; the matrix A has full rank.

Left inverse

Recall that A has full column rank if its columns are independent; i.e. if $r = n$. In this case the nullspace of A contains just the zero vector. The equation $A\mathbf{x} = \mathbf{b}$ either has exactly one solution $\mathbf{x}$ or is not solvable.

The matrix A^TA is an invertible n by n symmetric matrix, so $(A^TA)^{-1}A^TA = I$. We say $A_{\text{left}}^{-1} = (A^TA)^{-1}A^T$ is a *left inverse* of A. (There may be other left inverses as well, but this is our favorite.) The fact that A^TA is invertible when A has full column rank was central to our discussion of least squares.

Note that AA_{left}^{-1} is an m by m matrix which only equals the identity if $m = n$. A rectangular matrix can't have a two sided inverse because either that matrix or its transpose has a nonzero nullspace.

Right inverse

If A has full row rank, then $r = m$. The nullspace of A^T contains only the zero vector; the rows of A are independent. The equation $A\mathbf{x} = \mathbf{b}$ always has at least one solution; the nullspace of A has dimension $n - m$, so there will be $n - m$ free variables and (if $n > m$) infinitely many solutions!

Matrices with full row rank have right inverses A_{right}^{-1} with $AA_{\text{right}}^{-1} = I$. The nicest one of these is $A^T(AA^T)^{-1}$. Check: A times $A^T(AAT)^{-1}$ is I.

Pseudoinverse

An invertible matrix ($r = m = n$) has only the zero vector in its nullspace and left nullspace. A matrix with full column rank $r = n$ has only the zero vector in its nullspace. A matrix with full row rank $r = m$ has only the zero vector in its left nullspace. The remaining case to consider is a matrix A for which $r < n$ and $r < m$.

If A has full column rank and $A_{\text{left}}^{-1} = (A^TA)^{-1}A^T$, then

$$AA_{\text{left}}^{-1} = A(A^TA)^{-1}A^T = P$$

is the matrix which projects $\mathbb{R}^m$ onto the column space of A. This is as close as we can get to the product $AM = I$.

Similarly, if A has full row rank then $A_{\text{right}}^{-1}A = A^T(AA^T)^{-1}A$ is the matrix which projects $\mathbb{R}^n$ onto the row space of A.

It's nontrivial nullspaces that cause trouble when we try to invert matrices. If $A\mathbf{x} = \mathbf{0}$ for some nonzero $\mathbf{x}$, then there's no hope of finding a matrix A^{-1} that will reverse this process to give $A^{-1}\mathbf{0} = \mathbf{x}$.

The vector $A\mathbf{x}$ is always in the column space of A. In fact, the correspondence between vectors $\mathbf{x}$ in the (r dimensional) row space and vectors $A\mathbf{x}$ in the (r dimensional) column space is one-to-one. In other words, if $\mathbf{x} \neq \mathbf{y}$ are vectors in the row space of A then $A\mathbf{x} \neq A\mathbf{y}$ in the column space of A. (The proof of this would make a good exam question.)

Proof that if $\mathbf{x} \neq \mathbf{y}$ then $A\mathbf{x} \neq A\mathbf{y}$

Suppose the statement is false. Then we can find $\mathbf{x} \neq \mathbf{y}$ in the row space of A for which $A\mathbf{x} = A\mathbf{y}$. But then $A(\mathbf{x} - \mathbf{y}) = \mathbf{0}$, so $\mathbf{x} - \mathbf{y}$ is in the nullspace of A. But the row space of A is closed under linear combinations (like subtraction), so $\mathbf{x} - \mathbf{y}$ is also in the row space. The only vector in both the nullspace and the row space is the zero vector, so $\mathbf{x} - \mathbf{y} = \mathbf{0}$. This contradicts our assumption that $\mathbf{x}$ and $\mathbf{y}$ are not equal to each other.

We conclude that the mapping $\mathbf{x} \mapsto A\mathbf{x}$ from row space to column space is invertible. The inverse of this operation is called the *pseudoinverse* and is very useful to statisticians in their work with linear regression – they might not be able to guarantee that their matrices have full column rank $r = n$.

Finding the pseudoinverse A^+

The *pseudoinverse* A^+ of A is the matrix for which $\mathbf{x} = A^+A\mathbf{x}$ for all $\mathbf{x}$ in the row space of A. The nullspace of A^+ is the nullspace of A^T.

We start from the singular value decomposition $A = U\Sigma V^T$. Recall that Σ is a m by n matrix whose entries are zero except for the singular values $\sigma_1, \sigma_2, ..., \sigma_r$ which appear on the diagonal of its first r rows. The matrices U and V are orthonormal and therefore easy to invert. We only need to find a pseudoinverse for Σ.

The closest we can get to an inverse for Σ is an n by m matrix Σ^+ whose first r rows have $1/\sigma_1, 1/\sigma_2, ..., 1/\sigma_r$ on the diagonal. If $r = n = m$ then $\Sigma^+ = \Sigma^{-1}$. Always, the product of Σ and Σ^+ is a square matrix whose first r diagonal entries are 1 and whose other entries are 0.

If $A = U\Sigma V^T$ then its pseudoinverse is $A^+ = V\Sigma^+\mathbf{U}^T$. (Recall that $Q^T = Q^{-1}$ for orthogonal matrices U, V or Q.)

We would get a similar result if we included non-zero entries in the lower right corner of Σ^+, but we prefer not to have extra non-zero entries.

Conclusion

Although pseudoinverses will not appear on the exam, many of the topics we covered while discussing them (the four subspaces, the SVD, orthogonal matrices) are likely to appear.