Week III

Vector spaces

In this week's lectures, we learn about vector spaces. A vector space consists of a set of vectors and a set of scalars that is closed under vector addition and scalar multiplication and that satisfies the usual rules of arithmetic. We will learn some of the vocabulary and phrases of linear algebra, such as linear independence, span, basis and dimension. We will learn about the four fundamental subspaces of a matrix, the Gram-Schmidt process, orthogonal projection, and the matrix formulation of the least-squares problem of drawing a straight line to fit noisy data.

Lecture 16

Vector spaces

View this lecture on YouTube

A *vector space* consists of a set of vectors and a set of scalars. Although vectors can be quite general, for the purpose of this course we will only consider vectors that are real column matrices. The set of scalars can either be the real or complex numbers, and here we will only consider real numbers.

For the set of vectors and scalars to form a vector space, the set of vectors must be closed under vector addition and scalar multiplication. That is, when you multiply any two vectors in the set by real numbers and add them, the resulting vector must still be in the set.

As an example, consider the set of vectors consisting of all three-by-one column matrices, and let u and v be two of these vectors. Let $\mathrm{w} = a\mathrm{u} + b\mathrm{v}$ be the sum of these two vectors multiplied by the real numbers a and b. If w is still a three-by-one matrix, that is, w is in the set of vectors consisting of all three-by-one column matrices, then this set of vectors is closed under scalar multiplication and vector addition, and is indeed a vector space. The proof is rather simple. If we let

$$\mathrm{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \mathrm{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

then

$$\mathrm{w} = a\mathrm{u} + b\mathrm{v} = \begin{pmatrix} au_1 + bv_1 \\ au_2 + bv_2 \\ au_3 + bv_3 \end{pmatrix}$$

is evidently a three-by-one matrix, so that the set of all three-by-one matrices (together with the set of real numbers) is a vector space. This space is usually called $\mathbb{R}^3$.

Our main interest in vector spaces is to determine the vector spaces associated with matrices. There are four fundamental vector spaces of an m-by-n matrix A. They are called the *null space*, the *column space*, the *row space*, and the *left null space*. We will meet these vector spaces in later lectures.

Problems for Lecture 16

1. Explain why the zero vector must be a member of every vector space.

2. Explain why the following sets of three-by-one matrices (with real number scalars) are vector spaces:

(a) The set of three-by-one matrices with zero in the first row;

(b) The set of three-by-one matrices with first row equal to the second row;

(c) The set of three-by-one matrices with first row a constant multiple of the third row.

Solutions to the Problems

Lecture 17

Linear independence

View this lecture on YouTube

The set of vectors, $\{\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_n\}$, are *linearly independent* if for any scalars c_1, c_2, $\dots$, c_n, the equation

$$c_1\mathrm{u}_1 + c_2\mathrm{u}_2 + \cdots + c_n\mathrm{u}_n = 0$$

has only the solution $c_1 = c_2 = \cdots = c_n = 0$. What this means is that one is unable to write any of the vectors $\mathrm{u}_1, \mathrm{u}_2, \dots, \mathrm{u}_n$ as a linear combination of any of the other vectors. For instance, if there was a solution to the above equation with $c_1 \neq 0$, then we could solve that equation for u_1 in terms of the other vectors with nonzero coefficients.

As an example consider whether the following three three-by-one column vectors are linearly independent:

$$\mathrm{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathrm{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathrm{w} = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}.$$

Indeed, they are not linearly independent, that is, they are *linearly dependent*, because w can be written in terms of u and v. In fact, $\mathrm{w} = 2\mathrm{u} + 3\mathrm{v}$.

Now consider the three three-by-one column vectors given by

$$\mathrm{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathrm{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathrm{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

These three vectors are linearly independent because you cannot write any one of these vectors as a linear combination of the other two. If we go back to our definition of linear independence, we can see that the equation

$$a\mathrm{u} + b\mathrm{v} + c\mathrm{w} = \begin{pmatrix} a \\ b \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

has as its only solution $a = b = c = 0$.

Problems for Lecture 17

1. Which of the following sets of vectors are linearly independent?

(a) $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$

(b) $\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$

(c) $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

Solutions to the Problems

Lecture 18

Span, basis and dimension

View this lecture on YouTube

Given a set of vectors, one can generate a vector space by forming all linear combinations of that set of vectors. The *span* of the set of vectors $\{v_1, v_2, \ldots, v_n\}$ is the vector space consisting of all linear combinations of $v_1, v_2, \ldots, v_n$. We say that a set of vectors spans a vector space.

For example, the set of vectors given by

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}$$

spans the vector space of all three-by-one matrices with zero in the third row. This vector space is a *vector subspace* of all three-by-one matrices.

One doesn't need all three of these vectors to span this vector subspace because any one of these vectors is linearly dependent on the other two. The smallest set of vectors needed to span a vector space forms a *basis* for that vector space. Here, given the set of vectors above, we can construct a basis for the vector subspace of all three-by-one matrices with zero in the third row by simply choosing two out of three vectors from the above spanning set. Three possible bases are given by

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \right\}.$$

Although all three combinations form a basis for the vector subspace, the first combination is usually preferred because this is an orthonormal basis. The vectors in this basis are mutually orthogonal and of unit norm.

The number of vectors in a basis gives the dimension of the vector space. Here, the dimension of the vector space of all three-by-one matrices with zero in the third row is two.

Problems for Lecture 18

1. Find an orthonormal basis for the vector space of all three-by-one matrices with first row equal to second row. What is the dimension of this vector space?

Solutions to the Problems

Practice quiz: Vector space definitions

1. Which set of three-by-one matrices (with real number scalars) is not a vector space?

a) The set of three-by-one matrices with zero in the second row.

b) The set of three-by-one matrices with the sum of all the rows equal to one.

c) The set of three-by-one matrices with the first row equal to the third row.

d) The set of three-by-one matrices with the first row equal to the sum of the second and third rows.

2. Which of the following sets of vectors are linearly independent?

a) $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$

b) $\left\{ \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \\ -2 \end{pmatrix} \right\}$

c) $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$

d) $\left\{ \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\}$

3. Which of the following is an orthonormal basis for the vector space of all three-by-one matrices with the sum of all rows equal to zero?

a) $\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}$

b) $\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$

c) $\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$

d) $\left\{ \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\}$

Solutions to the Practice quiz

Lecture 19

Gram-Schmidt process

View this lecture on YouTube

Given any basis for a vector space, we can use an algorithm called the Gram-Schmidt process to construct an orthonormal basis for that space. Let the vectors $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n$ be a basis for some n-dimensional vector space. We will assume here that these vectors are column matrices, but this process also applies more generally.

We will construct an orthogonal basis $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n$, and then normalize each vector to obtain an orthonormal basis. First, define $\mathrm{u}_1 = \mathrm{v}_1$. To find the next orthogonal basis vector, define

$$\mathrm{u}_2 = \mathrm{v}_2 - \frac{(\mathrm{u}_1^{\mathrm{T}}\mathrm{v}_2)\mathrm{u}_1}{\mathrm{u}_1^{\mathrm{T}}\mathrm{u}_1}.$$

Observe that u_2 is equal to v_2 minus the component of v_2 that is parallel to u_1. By multiplying both sides of this equation with $\mathrm{u}_1^{\mathrm{T}}$, it is easy to see that $\mathrm{u}_1^{\mathrm{T}}\mathrm{u}_2 = 0$ so that these two vectors are orthogonal.

The next orthogonal vector in the new basis can be found from

$$\mathrm{u}_3 = \mathrm{v}_3 - \frac{(\mathrm{u}_1^{\mathrm{T}}\mathrm{v}_3)\mathrm{u}_1}{\mathrm{u}_1^{\mathrm{T}}\mathrm{u}_1} - \frac{(\mathrm{u}_2^{\mathrm{T}}\mathrm{v}_3)\mathrm{u}_2}{\mathrm{u}_2^{\mathrm{T}}\mathrm{u}_2}.$$

Here, u_3 is equal to v_3 minus the components of v_3 that are parallel to u_1 and u_2. We can continue in this fashion to construct n orthogonal basis vectors. These vectors can then be normalized via

$$\widehat{\mathrm{u}}_1 = \frac{\mathrm{u}_1}{(\mathrm{u}_1^{\mathrm{T}}\mathrm{u}_1)^{1/2}}, \quad \text{etc.}$$

Since u_k is a linear combination of $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_k$, the vector subspace spanned by the first k basis vectors of the original vector space is the same as the subspace spanned by the first k orthonormal vectors generated through the Gram-Schmidt process. We can write this result as

$$\mathrm{span}\{\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_k\} = \mathrm{span}\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_k\}.$$

Problems for Lecture 19

1. Suppose the four basis vectors $\{v_1, v_2, v_3, v_4\}$ are given, and one performs the Gram-Schmidt process on these vectors in order. Write down the equation to find the fourth orthogonal vector u_4. Do not normalize.

Solutions to the Problems

Lecture 20

Gram-Schmidt process example

View this lecture on YouTube

As an example of the Gram-Schmidt process, consider a subspace of three-by-one column matrices with the basis

$$\{v_1, v_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\},$$

and construct an orthonormal basis for this subspace. Let $u_1 = v_1$. Then u_2 is found from

$$\begin{aligned} u_2 &= v_2 - \frac{(u_1^T v_2)u_1}{u_1^T u_1} \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Normalizing the two vectors, we obtain the orthonormal basis

$$\{\hat{u}_1, \hat{u}_2\} = \left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{6}}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Notice that the initial two vectors v_1 and v_2 span the vector subspace of three-by-one column matrices for which the second and third rows are equal. Clearly, the orthonormal basis vectors constructed from the Gram-Schmidt process span the same subspace.

Problems for Lecture 20

1. Consider the vector subspace of three-by-one column vectors with the third row equal to the negative of the second row, and with the following given basis:

$$W = \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

Use the Gram-Schmidt process to construct an orthonormal basis for this subspace.

2. Consider a subspace of all four-by-one column vectors with the following basis:

$$W = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Use the Gram-Schmidt process to construct an orthonormal basis for this subspace.

Solutions to the Problems

Practice quiz: Gram-Schmidt process

1. In the fourth step of the Gram-Schmidt process, the vector $u_4 = v_4 - \frac{(u_1^T v_4)u_1}{u_1^T u_1} - \frac{(u_2^T v_4)u_2}{u_2^T u_2} - \frac{(u_3^T v_4)u_3}{u_3^T u_3}$ is always perpendicular to

a) v_1

b) v_2

c) v_3

d) v_4

2. The Gram-Schmidt process applied to $\{v_1, v_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ results in

a) $\{\hat{u}_1, \hat{u}_2\} = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$

b) $\{\hat{u}_1, \hat{u}_2\} = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$

c) $\{\hat{u}_1, \hat{u}_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$

d) $\{\hat{u}_1, \hat{u}_2\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\}$

3. The Gram-Schmidt process applied to $\{\mathrm{v}_1, \mathrm{v}_2\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$ results in

a) $\{\widehat{\mathrm{u}}_1, \widehat{\mathrm{u}}_2\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$

b) $\{\widehat{\mathrm{u}}_1, \widehat{\mathrm{u}}_2\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} \right\}$

c) $\{\widehat{\mathrm{u}}_1, \widehat{\mathrm{u}}_2\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$

d) $\{\widehat{\mathrm{u}}_1, \widehat{\mathrm{u}}_2\} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$

Solutions to the Practice quiz

Lecture 21

Null space

View this lecture on YouTube

The null space of a matrix A, which we denote as Null(A), is the vector space spanned by all column vectors x that satisfy the matrix equation

$$\mathrm{Ax} = 0.$$

Clearly, if x and y are in the null space of A, then so is $a\mathrm{x} + b\mathrm{y}$ so that the null space is closed under vector addition and scalar multiplication. If the matrix A is m-by-n, then Null(A) is a vector subspace of all n-by-one column matrices. If A is a square invertible matrix, then Null(A) consists of just the zero vector.

To find a basis for the null space of a noninvertible matrix, we bring A to reduced row echelon form. We demonstrate by example. Consider the three-by-five matrix given by

$$\mathrm{A} = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}.$$

By judiciously permuting rows to simplify the arithmetic, one pathway to construct rref(A) is

$$\begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ -3 & 6 & -1 & 1 & -7 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ 0 & 0 & 5 & 10 & -10 \\ 0 & 0 & 1 & 2 & -2 \end{pmatrix} \to$$
$$\begin{pmatrix} 1 & -2 & 2 & 3 & -1 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 5 & 10 & -10 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We call the variables associated with the pivot columns, x_1 and x_3, *basic variables*, and the variables associated with the non-pivot columns, x_2, x_4 and x_5, *free variables*. Writing the basic variables on the left-hand side of the $\mathrm{Ax} = 0$ equations, we have from the first and second rows

$$\begin{aligned} x_1 &= 2x_2 + x_4 - 3x_5, \\ x_3 &= -2x_4 + 2x_5. \end{aligned}$$

Eliminating x_1 and x_3, we can write the general solution for vectors in Null(A) as

$$\begin{pmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix},$$

where the free variables x_2, x_4, and x_5 can take any values. By writing the null space in this form, a basis for Null(A) is made evident, and is given by

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

The null space of A is seen to be a three-dimensional subspace of all five-by-one column matrices. In general, the dimension of Null(A) is equal to the number of non-pivot columns of rref(A).

Problems for Lecture 21

1. Determine a basis for the null space of

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

Solutions to the Problems

Lecture 22

Application of the null space

View this lecture on YouTube

An under-determined system of linear equations $\mathrm{A}x = \mathrm{b}$ with more unknowns than equations may not have a unique solution. If u is the general form of a vector in the null space of A, and v is any vector that satisfies $\mathrm{Av} = \mathrm{b}$, then $x = \mathrm{u} + \mathrm{v}$ satisfies $\mathrm{A}x = \mathrm{A}(\mathrm{u}+\mathrm{v}) = \mathrm{Au} + \mathrm{Av} = 0 + b = b$. The general solution of $\mathrm{A}x = \mathrm{b}$ can therefore be written as the sum of a general vector in $\mathrm{Null}(\mathrm{A})$ and a particular vector that satisfies the under-determined system.

As an example, suppose we want to find the general solution to the linear system of two equations and three unknowns given by

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &= 0, \\ 2x_1 - 2x_2 - x_3 &= 1, \end{aligned}$$

which in matrix form is given by

$$\begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We first bring the augmented matrix to reduced row echelon form:

$$\begin{pmatrix} 2 & 2 & 1 & 0 \\ 2 & -2 & -1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1/4 \\ 0 & 1 & 1/2 & -1/4 \end{pmatrix}.$$

The null space satisfying $\mathrm{Au} = 0$ is determined from $u_1 = 0$ and $u_2 = -u_3/2$, and we can write

$$\mathrm{Null}(\mathrm{A}) = \mathrm{span}\left\{ \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} \right\}.$$

A particular solution for the inhomogeneous system satisfying $\mathrm{Av} = b$ is found by solving $v_1 = 1/4$ and $v_2 + v_3/2 = -1/4$. Here, we simply take the free variable v_3 to be zero, and we find $v_1 = 1/4$ and $v_2 = -1/4$. The general solution to the original underdetermined linear system is the sum of the null space and the particular solution and is given by

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = a \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Problems for Lecture 22

1. Find the general solution to the system of equations given by

$$\begin{aligned} -3x_1 + 6x_2 - x_3 + x_4 &= -7, \\ x_1 - 2x_2 + 2x_3 + 3x_4 &= -1, \\ 2x_1 - 4x_2 + 5x_3 + 8x_4 &= -4. \end{aligned}$$

Solutions to the Problems

Lecture 23

Column space

View this lecture on YouTube

The column space of a matrix is the vector space spanned by the columns of the matrix. When a matrix is multiplied by a column vector, the resulting vector is in the column space of the matrix, as can be seen from

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix} = x \begin{pmatrix} a \\ c \end{pmatrix} + y \begin{pmatrix} b \\ d \end{pmatrix}.$$

In general, Ax is a linear combination of the columns of A. Given an m-by-n matrix A, what is the dimension of the column space of A, and how do we find a basis? Note that since A has m rows, the column space of A is a subspace of all m-by-one column matrices.

Fortunately, a basis for the column space of A can be found from rref(A). Consider the example

$$A = \begin{pmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{pmatrix}, \qquad \text{rref}(A) = \begin{pmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix equation $Ax = 0$ expresses the linear dependence of the columns of A, and row operations on A do not change the dependence relations. For example, the second column of A above is -2 times the first column, and after several row operations, the second column of rref(A) is still -2 times the first column.

It should be self-evident that only the pivot columns of rref(A) are linearly independent, and the dimension of the column space of A is therefore equal to its number of pivot columns; here it is two. A basis for the column space is given by the first and third columns of A, (not rref(A)), and is

$$\left\{ \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix} \right\}.$$

Recall that the dimension of the null space is the number of non-pivot columns—equal to the number of free variables—so that the sum of the dimensions of the null space and the column space is equal to the total number of columns. A statement of this theorem is as follows. Let A be an m-by-n matrix. Then

$$\dim(\text{Col}(A)) + \dim(\text{Null}(A)) = n.$$

Problems for Lecture 23

1. Determine the dimension and find a basis for the column space of

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

Solutions to the Problems

Lecture 24

Row space, left null space and rank

View this lecture on YouTube

In addition to the column space and the null space, a matrix A has two more vector spaces associated with it, namely the column space and null space of A^T, which are called the row space and the left null space.

If A is an m-by-n matrix, then the row space and the null space are subspaces of all n-by-one column matrices, and the column space and the left null space are subspaces of all m-by-one column matrices.

The null space consists of all vectors x such that $Ax = 0$, that is, the null space is the set of all vectors that are orthogonal to the row space of A. We say that these two vector spaces are orthogonal.

A basis for the row space of a matrix can be found from computing rref(A), and is found to be rows of rref(A) (written as column vectors) with pivot columns. The dimension of the row space of A is therefore equal to the number of pivot columns, while the dimension of the null space of A is equal to the number of nonpivot columns. The union of these two subspaces make up the vector space of all n-by-one matrices and we say that these subspaces are *orthogonal complements* of each other.

Furthermore, the dimension of the column space of A is also equal to the number of pivot columns, so that the dimensions of the column space and the row space of a matrix are equal. We have

$$\dim(\text{Col}(A)) = \dim(\text{Row}(A)).$$

We call this dimension the rank of the matrix A. This is an amazing result since the column space and row space are subspaces of two different vector spaces. In general, we must have $\text{rank}(A) \leq \min(m, n)$. When the equality holds, we say that the matrix is of full rank. And when A is a square matrix and of full rank, then the dimension of the null space is zero and A is invertible.

Problems for Lecture 24

1. Find a basis for the column space, row space, null space and left null space of the four-by-five matrix A, where

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

Check to see that null space is the orthogonal complement of the row space, and the left null space is the orthogonal complement of the column space. Find $\mathrm{rank}(A)$. Is this matrix of full rank?

Solutions to the Problems

Practice quiz: Fundamental subspaces

1. Which of the following sets of vectors form a basis for the null space of $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 1 & 1 \\ 3 & 6 & 1 & 1 \end{pmatrix}$?

a) $\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \\ 0 \\ 0 \end{pmatrix} \right\}$

b) $\left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

c) $\left\{ \begin{pmatrix} 0 \\ 0 \\ -3 \\ 2 \end{pmatrix} \right\}$

d) $\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$

2. The general solution to the system of equations given by

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 1, \\ 2x_1 + 4x_2 + x_3 + x_4 &= 1, \\ 3x_1 + 6x_2 + x_3 + x_4 &= 1, \end{aligned}$$

is

a) $a \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

b) $a \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

c) $a \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -3 \\ 2 \end{pmatrix}$

d) $a \begin{pmatrix} 0 \\ 0 \\ -3 \\ 2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

3. What is the rank of the matrix $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 4 & 1 & 1 \\ 3 & 6 & 1 & 1 \end{pmatrix}$?

a) 1

b) 2

c) 3

d) 4

Solutions to the Practice quiz

Lecture 25

Orthogonal projections

View this lecture on YouTube

Suppose that V is an n-dimensional vector space and W is a p-dimensional subspace of V. For familiarity, we assume here that all vectors are column matrices of fixed size. Let v be a vector in V and let $\{\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_p\}$ be an orthonormal basis for W. In general, the orthogonal projection of a vector v in V onto the subspace W is given by

$$\mathrm{v}_{\mathrm{proj}_W} = (\mathrm{v}^{\mathrm{T}}\mathrm{s}_1)\mathrm{s}_1 + (\mathrm{v}^{\mathrm{T}}\mathrm{s}_2)\mathrm{s}_2 + \cdots + (\mathrm{v}^{\mathrm{T}}\mathrm{s}_p)\mathrm{s}_p;$$

and by adding and subtracting $\mathrm{v}_{\mathrm{proj}_W}$, we can write

$$\mathrm{v} = \mathrm{v}_{\mathrm{proj}_W} + (\mathrm{v} - \mathrm{v}_{\mathrm{proj}_W}),$$

where $\mathrm{v}_{\mathrm{proj}_W}$ is a vector in W and $(\mathrm{v} - \mathrm{v}_{\mathrm{proj}_W})$ is a vector orthogonal to W.

We can be more concrete. Using the Gram-Schmidt process, it is possible to construct a basis for the vector space V consisting of all the orthonormal basis vectors for the subspace W together with whatever remaining orthonormal vectors are required to span V. Write this basis for the n-dimensional vector space V as $\{\mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_p, \mathrm{t}_1, \mathrm{t}_2, \ldots, \mathrm{t}_{n-p}\}$. Then any vector v in V can be written as

$$\mathrm{v} = a_1\mathrm{s}_1 + a_2\mathrm{s}_2 + \cdots + a_p\mathrm{s}_p + b_1\mathrm{t}_1 + b_2\mathrm{t}_2 + b_{n-p}\mathrm{t}_{n-p}.$$

Using the orthonormality of the basis vectors, the orthogonal projection of v onto W in then seen to be

$$\mathrm{v}_{\mathrm{proj}_W} = a_1\mathrm{s}_1 + a_2\mathrm{s}_2 + \cdots + a_p\mathrm{s}_p,$$

that is, the part of v that lies in W.

The vector $\mathrm{v}_{\mathrm{proj}_W}$ is the vector in W that is closest to v. Let w be any vector in W different than $\mathrm{v}_{\mathrm{proj}_W}$, and expand w in terms of the basis vectors for W:

$$\mathrm{w} = c_1\mathrm{s}_1 + c_2\mathrm{s}_2 + \cdots + c_p\mathrm{s}_p.$$

The distance between v and w is given by the norm $||\mathrm{v} - \mathrm{w}||$, and we have

$$\begin{aligned} ||\mathrm{v} - \mathrm{w}||^2 &= (a_1 - c_1)^2 + (a_2 - c_2)^2 + \cdots + (a_p - c_p)^2 + b_1^2 + b_2^2 + \cdots + b_{n-p}^2 \\ &\geq b_1^2 + b_2^2 + \cdots + b_{n-p}^2 = ||\mathrm{v} - \mathrm{v}_{\mathrm{proj}_W}||^2, \end{aligned}$$

or $||\mathrm{v} - \mathrm{v}_{\mathrm{proj}_W}|| \leq ||\mathrm{v} - \mathrm{w}||$, a result that will be used later in the problem of least squares.

Problems for Lecture 25

1. Find the general orthogonal projection of $\mathbf{v}$ onto W, where $\mathbf{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ and $W = \text{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$. What are the projections when $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and when $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$?

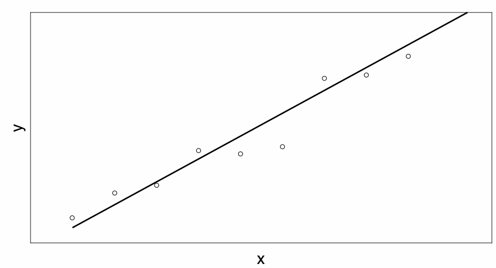
Solutions to the Problems

Lecture 26

The least-squares problem

View this lecture on YouTube

Suppose there is some experimental data that you want to fit by a straight line. This is called a linear regression problem and an illustrative example is shown below.



Linear regression

In general, let the data consist of a set of n points given by (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n). Here, we assume that the x values are exact, and the y values are noisy. We further assume that the best fit line to the data takes the form $y = \beta_0 + \beta_1 x$. Although we know that the line will not go through all of the data points, we can still write down the equations as if it does. We have

$$y_1 = \beta_0 + \beta_1 x_1, \quad y_2 = \beta_0 + \beta_1 x_2, \quad \ldots, \quad y_n = \beta_0 + \beta_1 x_n.$$

These equations constitute a system of n equations in the two unknowns β_0 and β_1. The corresponding matrix equation is given by

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

This is an overdetermined system of equations with no solution. The problem of least squares is to find the best solution.

We can generalize this problem as follows. Suppose we are given a matrix equation, $\mathrm{Ax} = \mathrm{b}$, that has no solution because b is not in the column space of A. So instead we solve $\mathrm{Ax} = \mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}}$, where $\mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}}$ is the projection of b onto the column space of A. The solution is then called the least-squares solution for x.

Problems for Lecture 26

1. Suppose we have data points given by $(x_i, y_i) = (0, 1)$, $(1, 3)$, $(2, 3)$, and $(3, 4)$. If the data is to be fit by the line $y = \beta_0 + \beta_1 x$, write down the overdetermined matrix expression for the set of equations $y_i = \beta_0 + \beta_1 x_i$.

Solutions to the Problems

Lecture 27

Solution of the least-squares problem

View this lecture on YouTube

We want to find the least-squares solution to an overdetermined matrix equation $Ax = b$. We write $b = b_{\text{proj}_{\text{Col}(A)}} + (b - b_{\text{proj}_{\text{Col}(A)}})$, where $b_{\text{proj}_{\text{Col}(A)}}$ is the projection of b onto the column space of A. Since $(b - b_{\text{proj}_{\text{Col}(A)}})$ is orthogonal to the column space of A, it is in the nullspace of A^T. Multiplication of the overdetermined matrix equation by A^T then results in a solvable set of equations, called the *normal equations* for $Ax = b$, given by

$$A^T A x = A^T b.$$

A unique solution to this matrix equation exists when the columns of A are linearly independent.

An interesting formula exists for the matrix which projects b onto the column space of A. Multiplying the normal equations on the left by $A(A^T A)^{-1}$, we obtain

$$Ax = A(A^T A)^{-1} A^T b = b_{\text{proj}_{\text{Col}(A)}}.$$

Notice that the projection matrix $P = A(A^T A)^{-1} A^T$ satisfies $P^2 = P$, that is, two projections is the same as one. If A itself is a square invertible matrix, then $P = I$ and b is already in the column space of A.

As an example of the application of the normal equations, consider the toy least-squares problem of fitting a line through the three data points $(1, 1)$, $(2, 3)$ and $(3, 2)$. With the line given by $y = \beta_0 + \beta_1 x$, the overdetermined system of equations is given by

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}.$$

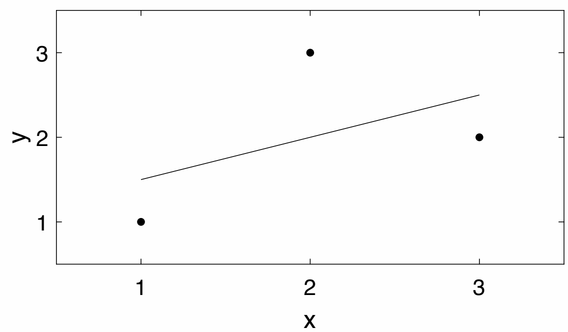
The least-squares solution is determined by solving

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix},$$

or

$$\begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 6 \\ 13 \end{pmatrix}.$$

We can using Gaussian elimination to determine $\beta_0 = 1$ and $\beta_1 = 1/2$, and the least-squares line is given by $y = 1 + x/2$. The graph of the data and the line is shown below.



Solution of a toy least-squares problem.

Problems for Lecture 27

1. Suppose we have data points given by $(x_n, y_n) = (0, 1)$, $(1, 3)$, $(2, 3)$, and $(3, 4)$. By solving the normal equations, fit the data by the line $y = \beta_0 + \beta_1 x$.

Solutions to the Problems

Practice quiz: Orthogonal projections

1. Which vector is the orthogonal projection of $\mathrm{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ onto $W = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$?

a) $\frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$

b) $\frac{1}{3}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$

c) $\frac{1}{3}\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$

d) $\frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

2. Suppose we have data points given by $(x_n, y_n) = (1,1)$, $(2,1)$, and $(3,3)$. If the data is to be fit by the line $y = \beta_0 + \beta_1 x$, which is the overdetermined equation for β_0 and β_1?

a) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$

c) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$

3. Suppose we have data points given by $(x_n, y_n) = (1,1)$, $(2,1)$, and $(3,3)$. Which is the best fit line to the data?

a) $y = \frac{1}{3} + x$

b) $y = -\frac{1}{3} + x$

c) $y = 1 + \frac{1}{3}x$

d) $y = 1 - \frac{1}{3}x$

Solutions to the Practice quiz