Matrix Algebra for Engineers

Jeffrey R. Chasnov

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

The Hong Kong University of
Science and Technology

The Hong Kong University of Science and Technology
Department of Mathematics
Clear Water Bay, Kowloon
Hong Kong

Preface

These are my lecture notes for my online Coursera course, Matrix Algebra for Engineers. I have divided these notes into chapters called Lectures, with each Lecture corresponding to a video on Coursera. I have also uploaded all my Coursera videos to YouTube, and links are placed at the top of each Lecture.

There are problems at the end of each lecture chapter and I have tried to choose problems that exemplify the main idea of the lecture. Students taking a formal university course in matrix or linear algebra will usually be assigned many more additional problems, but here I follow the philosophy that less is more. I give enough problems for students to solidify their understanding of the material, but not too many problems that students feel overwhelmed and drop out. I do encourage students to attempt the given problems, but if they get stuck, full solutions can be found in the Appendix.

The mathematics in this matrix algebra course is at the level of an advanced high school student, but typically students would take this course after completing a university-level single variable calculus course. There are no derivatives and integrals in this course, but student's are expected to have a certain level of mathematical maturity. Nevertheless, anyone who wants to learn the basics of matrix algebra is welcome to join.

Jeffrey R. Chasnov
Hong Kong
July 2018

Contents

Week I

Matrices

In this week's lectures, we learn about matrices. Matrices are rectangular arrays of numbers or other mathematical objects and are fundamental to engineering mathematics. We will define matrices and how to add and multiply them, discuss some special matrices such as the identity and zero matrix, learn about transposes and inverses, and define orthogonal and permutation matrices.

Lecture 1

Definition of a matrix

View this lecture on YouTube

An m-by-n matrix is a rectangular array of numbers (or other mathematical objects) with m rows and n columns. For example, a two-by-two matrix A, with two rows and two columns, looks like

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The first row has elements a and b, the second row has elements c and d. The first column has elements a and c; the second column has elements b and d. As further examples, two-by-three and three-by-two matrices look like

$$\mathrm{B} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \quad \mathrm{C} = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}.$$

Of special importance are column matrices and row matrices. These matrices are also called vectors. The column vector is in general n-by-one and the row vector is one-by-n. For example, when $n = 3$, we would write a column vector as

$$\mathrm{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix},$$

and a row vector as

$$\mathrm{y} = \begin{pmatrix} a & b & c \end{pmatrix}.$$

A useful notation for writing a general m-by-n matrix A is

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

Here, the matrix element of A in the ith row and the jth column is denoted as a_{ij}.

Problems for Lecture 1

1. The main diagonal of a matrix A are the entries a_{ij} where $i = j$.

(a) Write down the three-by-three matrix with ones on the diagonal and zeros elsewhere.

(b) Write down the three-by-four matrix with ones on the diagonal and zeros elsewhere.

(c) Write down the four-by-three matrix with ones on the diagonal and zeros elsewhere.

Solutions to the Problems

Lecture 2

Addition and multiplication of matrices

View this lecture on YouTube

Matrices can be added only if they have the same dimension. Addition proceeds element by element. For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}.$$

Matrices can also be multiplied by a scalar. The rule is to just multiply every element of the matrix. For example,

$$k\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ka & kb \\ kc & kd \end{pmatrix}.$$

Matrices (other than the scalar) can be multiplied only if the number of columns of the left matrix equals the number of rows of the right matrix. In other words, an m-by-n matrix on the left can only be multiplied by an n-by-k matrix on the right. The resulting matrix will be m-by-k. Evidently, matrix multiplication is generally not commutative. We illustrate multiplication using two 2-by-2 matrices:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}, \qquad \begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ae+cf & be+df \\ ag+ch & bg+dh \end{pmatrix}.$$

First, the first row of the left matrix is multiplied against and summed with the first column of the right matrix to obtain the element in the first row and first column of the product matrix. Second, the first row is multiplied against and summed with the second column. Third, the second row is multiplied against and summed with the first column. And fourth, the second row is multiplied against and summed with the second column.

In general, an element in the resulting product matrix, say in row i and column j, is obtained by multiplying and summing the elements in row i of the left matrix with the elements in column j of the right matrix. We can formally write matrix multiplication in terms of the matrix elements. Let A be an m-by-n matrix with matrix elements a_{ij} and let B be an n-by-p matrix with matrix elements b_{ij}. Then $C = AB$ is an m-by-p matrix, and its ij matrix element can be written as

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Notice that the second index of a and the first index of b are summed over.

Problems for Lecture 2

1. Define the matrices

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -2 & 1 \\ 2 & -4 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$
$$D = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}, \quad E = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Compute if defined: $B - 2A$, $3C - E$, AC, CD, CB.

2. Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ and $C = \begin{pmatrix} 4 & 3 \\ 0 & 2 \end{pmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.

3. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{pmatrix}$ and $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$. Compute AD and DA.

Solutions to the Problems

Lecture 3

Special matrices

View this lecture on YouTube

The zero matrix, denoted by 0, can be any size and is a matrix consisting of all zero elements. Multiplication by a zero matrix results in a zero matrix. The identity matrix, denoted by I, is a square matrix (number of rows equals number of columns) with ones down the main diagonal. If A and I are the same sized square matrices, then

$$\mathrm{AI} = \mathrm{IA} = \mathrm{A},$$

and multiplication by the identity matrix leaves the matrix unchanged. The zero and identity matrices play the role of the numbers zero and one in matrix multiplication. For example, the two-by-two zero and identity matrices are given by

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A diagonal matrix has its only nonzero elements on the diagonal. For example, a two-by-two diagonal matrix is given by

$$\mathrm{D} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}.$$

Usually, diagonal matrices refer to square matrices, but they can also be rectangular.

A band (or banded) matrix has nonzero elements only on diagonal bands. For example, a three-by-three band matrix with nonzero diagonals one above and one below a nonzero main diagonal (called a tridiagonal matrix) is given by

$$\mathrm{B} = \begin{pmatrix} d_1 & a_1 & 0 \\ b_1 & d_2 & a_2 \\ 0 & b_2 & d_3 \end{pmatrix}.$$

An upper or lower triangular matrix is a square matrix that has zero elements below or above the diagonal. For example, three-by-three upper and lower triangular matrices are given by

$$\mathrm{U} = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}, \quad \mathrm{L} = \begin{pmatrix} a & 0 & 0 \\ b & d & 0 \\ c & e & f \end{pmatrix}.$$

Problems for Lecture 3

1. Let $A = \begin{pmatrix} -1 & 2 \\ 4 & -8 \end{pmatrix}$. Construct a two-by-two matrix B such that AB is the zero matrix. Use two different nonzero columns for B.

2. Verify that $\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & 0 \\ 0 & a_2 b_2 \end{pmatrix}$. Prove in general that the product of two diagonal matrices is a diagonal matrix, with elements given by the product of the diagonal elements.

3. Verify that $\begin{pmatrix} a_1 & a_2 \\ 0 & a_3 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ 0 & b_3 \end{pmatrix} = \begin{pmatrix} a_1 b_1 & a_1 b_2 + a_2 b_3 \\ 0 & a_3 b_3 \end{pmatrix}$. Prove in general that the product of two upper triangular matrices is an upper triangular matrix, with the diagonal elements of the product given by the product of the diagonal elements.

Solutions to the Problems

Lecture 4

Transpose matrix

View this lecture on YouTube

The transpose of a matrix A, denoted by A^T and spoken as A-transpose, switches the rows and columns of A. That is,

$$\text{if } A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \text{then } A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

In other words, we write

$$a_{ij}^T = a_{ji}.$$

Evidently, if A is m-by-n then A^T is n-by-m. As a simple example, view the following transpose pair:

$$\begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}^T = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

The following are useful and easy to prove facts:

$$\left(A^T\right)^T = A, \text{ and } (A+B)^T = A^T + B^T.$$

A less obvious fact is that the transpose of the product of matrices is equal to the product of the transposes with the order of multiplication reversed, i.e.,

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

If A is a square matrix, and $A^T = A$, then we say that A is *symmetric*. If $A^T = -A$, then we say that A is *skew symmetric*. For example, 3-by-3 symmetric and skew symmetric matrices look like

$$\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}, \quad \begin{pmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{pmatrix}.$$

Notice that the diagonal elements of a skew-symmetric matrix must be zero.

Problems for Lecture 4

1. Prove that $(AB)^T = B^T A^T$.

2. Show using the transpose operator that any square matrix A can be written as the sum of a symmetric and a skew-symmetric matrix.

3. Prove that $A^T A$ is symmetric.

Solutions to the Problems

Lecture 5

Inner and outer products

View this lecture on YouTube

The *inner product* (or dot product or scalar product) between two vectors is obtained from the matrix product of a row vector times a column vector. A row vector can be obtained from a column vector by the transpose operator. With the 3-by-1 column vectors u and v, their inner product is given by

$$\mathrm{u}^{\mathrm{T}}\mathrm{v} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = u_1v_1 + u_2v_2 + u_3v_3.$$

If the inner product between two vectors is zero, we say that the vectors are *orthogonal*. The *norm* of a vector is defined by

$$||\mathrm{u}|| = \left(\mathrm{u}^{\mathrm{T}}\mathrm{u}\right)^{1/2} = \left(u_1^2 + u_2^2 + u_3^2\right)^{1/2}.$$

If the norm of a vector is equal to one, we say that the vector is *normalized*. If a set of vectors are mutually orthogonal and normalized, we say that these vectors are *orthonormal*.

An *outer product* is also defined, and is used in some applications. The outer product between u and v is given by

$$\mathrm{u}\mathrm{v}^{\mathrm{T}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} = \begin{pmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{pmatrix}.$$

Notice that every column is a multiple of the single vector u, and every row is a multiple of the single vector v^{T}.

Problems for Lecture 5

1. Let A be a rectangular matrix given by $\mathrm{A} = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}$. Compute $\mathrm{A}^{\mathrm{T}}\mathrm{A}$ and show that it is a symmetric square matrix and that the sum of its diagonal elements is the sum of the squares of all the elements of A.

2. The trace of a square matrix B, denoted as $\operatorname{Tr}\mathrm{B}$, is the sum of the diagonal elements of B. Prove that $\operatorname{Tr}(\mathrm{A}^{\mathrm{T}}\mathrm{A})$ is the sum of the squares of all the elements of A.

Solutions to the Problems

Lecture 6

Inverse matrix

View this lecture on YouTube

Square matrices may have inverses. When a matrix A has an inverse, we say it is invertible and denote its inverse by A^{-1}. The inverse matrix satisfies

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

If A and B are same-sized square matrices, and $AB = I$, then $A = B^{-1}$ and $B = A^{-1}$. In words, the right and left inverses of square matrices are equal. Also, $(AB)^{-1} = B^{-1}A^{-1}$. In words, the inverse of the product of invertible matrices is equal to the product of the inverses with the order of multiplication reversed. Finally, if A is invertible then so is A^T, and $(A^T)^{-1} = (A^{-1})^T$. In words, the inverse of the transpose matrix is the transpose of the inverse matrix.

It is illuminating to derive the inverse of a general 2-by-2 matrix. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

and write

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We solve for x_1, y_1, x_2 and y_2. There are two inhomogeneous and two homogeneous linear equations:

$$ax_1 + by_1 = 1, \qquad cx_1 + dy_1 = 0,$$
$$cx_2 + dy_2 = 1, \qquad ax_2 + by_2 = 0.$$

To solve, we can eliminate y_1 and y_2 using the two homogeneous equations, and then solve for x_1 and x_2 using the two inhomogeneous equations. Finally, we use the two homogeneous equations to solve for y_1 and y_2. The solution for A^{-1} is found to be

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

The factor in front of the matrix is in fact the definition of the determinant of a two-by-two matrix A:

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

The determinant of a two-by-two matrix is the product of the diagonals minus the product of the off-diagonals. Evidently, A is invertible only if $\det A \neq 0$. Notice that the inverse of a two-by-two matrix, in words, is found by switching the diagonal elements of the matrix, negating the off-diagonal elements, and dividing by the determinant.

Later, we will show that an n-by-n matrix is invertible if and only if its determinant is nonzero. This will require a more general definition of determinant.

Problems for Lecture 6

1. Find the inverses of the matrices $\begin{pmatrix} 5 & 6 \\ 4 & 5 \end{pmatrix}$ and $\begin{pmatrix} 6 & 4 \\ 3 & 3 \end{pmatrix}$.

2. Prove that if A and B are same-sized invertible matrices , then $(AB)^{-1} = B^{-1}A^{-1}$.

3. Prove that if A is invertible then so is A^T, and $(A^T)^{-1} = (A^{-1})^T$.

4. Prove that if a matrix is invertible, then its inverse is unique.

Solutions to the Problems

Lecture 7

Orthogonal matrices

View this lecture on YouTube

A square matrix Q with real entries that satisfies

$$Q^{-1} = Q^{T}$$

is called an *orthogonal matrix*.

Since the columns of Q^T are just the rows of Q, and $QQ^T = I$, the row vectors that form Q must be orthonormal. Similarly, since the rows of Q^T are just the columns of Q, and $Q^TQ = I$, the column vectors that form Q must also be orthonormal.

Orthogonal matrices preserve norms. Let Q be an n-by-n orthogonal matrix, and let x be an n-by-one column vector. Then the norm squared of Qx is given by

$$||Qx||^2 = (Qx)^T (Qx) = x^TQ^TQx = x^Tx = ||x||^2.$$

The norm of a vector is also called its length, so we can also say that orthogonal matrices preserve lengths.

Problems for Lecture 7

1. Show that the product of two orthogonal matrices is orthogonal.

2. Show that the n-by-n identity matrix is orthogonal.

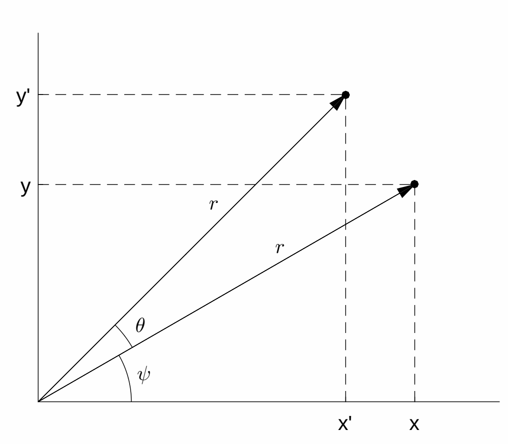
Solutions to the Problems

Lecture 8

Orthogonal matrices example

View this lecture on YouTube

A matrix that rotates a vector in space doesn't change the vector's length and so should be an orthog-



Rotating a vector in the x-y plane.

onal matrix. Consider the two-by-two rotation matrix that rotates a vector through an angle θ in the x-y plane, shown above. Trigonometry and the addition formula for cosine and sine results in

$$\begin{aligned} x' &= r\cos(\theta+\psi) \\ &= r(\cos\theta\cos\psi - \sin\theta\sin\psi) \\ &= x\cos\theta - y\sin\theta \end{aligned} \qquad \begin{aligned} y' &= r\sin(\theta+\psi) \\ &= r(\sin\theta\cos\psi + \cos\theta\sin\psi) \\ &= x\sin\theta + y\cos\theta. \end{aligned}$$

Writing the equations for x' and y' in matrix form, we have

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

The above two-by-two matrix is a rotation matrix and we will denote it by R_θ. Observe that the rows and columns of R_θ are orthonormal and that the inverse of R_θ is just its transpose. The inverse of R_θ rotates a vector by $-\theta$.

Problems for Lecture 8

1. Let $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Show that $R(-\theta) = R(\theta)^{-1}$.

2. Find the three-by-three matrix that rotates a three-dimensional vector an angle θ counterclockwise around the z-axis.

Solutions to the Problems

Lecture 9

Permutation matrices

View this lecture on YouTube

Another type of orthogonal matrix is a permutation matrix. An n-by-n permutation matrix, when multiplying on the left permutes the rows of a matrix, and when multiplying on the right permutes the columns. Clearly, permuting the rows of a column vector will not change its norm.

For example, let the string $\{1,2\}$ represent the order of the rows or columns of a two-by-two matrix. Then the permutations of the rows or columns are given by $\{1,2\}$ and $\{2,1\}$. The first permutation is no permutation at all, and the corresponding permutation matrix is simply the identity matrix. The second permutation of the rows or columns is achieved by

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}.$$

The rows or columns of a three-by-three matrix have $3! = 6$ possible permutations, namely $\{1,2,3\}$, $\{1,3,2\}$, $\{2,1,3\}$, $\{2,3,1\}$, $\{3,1,2\}$, $\{3,2,1\}$. For example, the row or column permutation $\{3,1,2\}$ is obtained by

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} g & h & i \\ a & b & c \\ d & e & f \end{pmatrix}, \qquad \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} c & a & b \\ f & d & e \\ i & g & h \end{pmatrix}.$$

Notice that the permutation matrix is obtained by permuting the corresponding rows (or columns) of the identity matrix. This is made evident by observing that

$$\mathrm{PA} = (\mathrm{PI})\mathrm{A}, \qquad \mathrm{AP} = \mathrm{A}(\mathrm{PI}),$$

where P is a permutation matrix and PI is the identity matrix with permuted rows. The identity matrix is orthogonal, and so is the matrix obtained by permuting its rows.

Problems for Lecture 9

1. Write down the six three-by-three permutation matrices corresponding to the permutations $\{1, 2, 3\}$, $\{1, 3, 2\}$, $\{2, 1, 3\}$, $\{2, 3, 1\}$, $\{3, 1, 2\}$, $\{3, 2, 1\}$.

2. Find the inverses of all the three-by-three permutation matrices. Explain why some matrices are their own inverses, and others are not.

Solutions to the Problems

Week II

Systems of linear equations

In this week's lectures, we learn about solving a system of linear equations. A system of linear equations can be written in matrix form, and we can solve using Gaussian elimination. We will learn how to bring a matrix to reduced row echelon form, and how this can be used to compute a matrix inverse. We will also learn how to find the LU decomposition of a matrix, and how to use this decomposition to efficiently solve a system of linear equations.

Lecture 10

Gaussian elimination

View this lecture on YouTube

Consider the linear system of equations given by

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

which can be written in matrix form as

$$\begin{pmatrix} -3 & 2 & -1 \\ 6 & -6 & 7 \\ 3 & -4 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -7 \\ -6 \end{pmatrix},$$

or symbolically as $\mathrm{Ax} = \mathrm{b}$.

The standard numerical algorithm used to solve a system of linear equations is called *Gaussian elimination*. We first form what is called an *augmented matrix* by combining the matrix A with the column vector b:

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 6 & -6 & 7 & -7 \\ 3 & -4 & 4 & -6 \end{pmatrix}.$$

Row reduction is then performed on this augmented matrix. Allowed operations are (1) interchange the order of any rows, (2) multiply any row by a constant, (3) add a multiple of one row to another row. These three operations do not change the solution of the original equations. The goal here is to convert the matrix A into upper-triangular form, and then use this form to quickly solve for the unknowns x.

We start with the first row of the matrix and work our way down as follows. First we multiply the first row by 2 and add it to the second row. Then we add the first row to the third row, to obtain

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & -2 & 3 & -7 \end{pmatrix}.$$

We then go to the second row. We multiply this row by -1 and add it to the third row to obtain

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & 0 & -2 & 2 \end{pmatrix}.$$

The original matrix A has been converted to an upper triangular matrix, and the transformed equations can be determined from the augmented matrix as

$$\begin{aligned} -3x_1 + 2x_2 - x_3 &= -1, \\ -2x_2 + 5x_3 &= -9, \\ -2x_3 &= 2. \end{aligned}$$

These equations can be solved by back substitution, starting from the last equation and working backwards. We have

$$\begin{aligned} x_3 &= -1, \\ x_2 &= -\frac{1}{2}(-9 - 5x_3) = 2, \\ x_1 &= -\frac{1}{3}(-1 + x_3 - 2x_2) = 2. \end{aligned}$$

We have thus found the solution

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

When performing Gaussian elimination, the diagonal element that is used during the elimination procedure is called the *pivot*. To obtain the correct multiple, one uses the pivot as the divisor to the matrix elements below the pivot. Gaussian elimination in the way done here will fail if the pivot is zero. If the pivot is zero, a row interchange must first be performed.

Even if no pivots are identically zero, small values can still result in an unstable numerical computation. For very large matrices solved by a computer, the solution vector will be inaccurate unless row interchanges are made. The resulting numerical technique is called Gaussian elimination with partial pivoting, and is usually taught in a standard numerical analysis course.

Problems for Lecture 10

1. Using Gaussian elimination with back substitution, solve the following two systems of equations:

(a)

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

(b)

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

Solutions to the Problems

Lecture 11

Reduced row echelon form

View this lecture on YouTube

If we continue the row elimination procedure so that all the pivots are one, and all the entries above and below the pivots are eliminated, then we say that the resulting matrix is in reduced row echelon form. We notate the reduced row echelon form of a matrix A as $\text{rref}(A)$. For example, consider the three-by-four matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{pmatrix}.$$

Row elimination can proceed as

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 \\ 6 & 7 & 8 & 9 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -6 & -9 \\ 0 & -5 & -10 & -15 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

and we therefore have

$$\text{rref}(A) = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We say that the matrix A has two pivot columns, that is two columns that contain a pivot position with a one in the reduced row echelon form. Note that rows may need to be exchanged when computing the reduced row echelon form.

Problems for Lecture 11

1. Put the following matrices into reduced row echelon form and state which columns are pivot columns:

(a)

$$A = \begin{pmatrix} 3 & -7 & -2 & -7 \\ -3 & 5 & 1 & 5 \\ 6 & -4 & 0 & 2 \end{pmatrix}$$

(b)

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

Solutions to the Problems

Lecture 12

Computing inverses

View this lecture on YouTube

By bringing an invertible matrix to reduced row echelon form, that is, to the identity matrix, we can compute the matrix inverse. Given a matrix A, consider the equation

$$\mathrm{A}\mathrm{A}^{-1} = \mathrm{I},$$

for the unknown inverse A^{-1}. Let the columns of A^{-1} be given by the vectors a_1^{-1}, a_2^{-1}, and so on. The matrix A multiplying the first column of A^{-1} is the equation

$$\mathrm{A}a_1^{-1} = \mathrm{e}_1, \quad \text{with } \mathrm{e}_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix}^{\mathrm{T}},$$

and where e_1 is the first column of the identity matrix. In general,

$$\mathrm{A}a_i^{-1} = \mathrm{e}_i,$$

for $i = 1, 2, \dots, n$. The method then is to do row reduction on an augmented matrix which attaches the identity matrix to A. To find A^{-1}, elimination is continued until one obtains $\mathrm{rref}(\mathrm{A}) = \mathrm{I}$.

We illustrate below:

$$\begin{pmatrix} -3 & 2 & -1 & 1 & 0 & 0 \\ 6 & -6 & 7 & 0 & 1 & 0 \\ 3 & -4 & 4 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} -3 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 5 & 2 & 1 & 0 \\ 0 & -2 & 3 & 1 & 0 & 1 \end{pmatrix} \rightarrow$$
$$\begin{pmatrix} -3 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 5 & 2 & 1 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} -3 & 0 & 4 & 3 & 1 & 0 \\ 0 & -2 & 5 & 2 & 1 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{pmatrix} \rightarrow$$
$$\begin{pmatrix} -3 & 0 & 0 & 1 & -1 & 2 \\ 0 & -2 & 0 & -1/2 & -3/2 & 5/2 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & -1/3 & 1/3 & -2/3 \\ 0 & 1 & 0 & 1/4 & 3/4 & -5/4 \\ 0 & 0 & 1 & 1/2 & 1/2 & -1/2 \end{pmatrix};$$

and one can check that

$$\begin{pmatrix} -3 & 2 & -1 \\ 6 & -6 & 7 \\ 3 & -4 & 4 \end{pmatrix} \begin{pmatrix} -1/3 & 1/3 & -2/3 \\ 1/4 & 3/4 & -5/4 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Problems for Lecture 12

1. Compute the inverse of

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

Solutions to the Problems

Lecture 13

Elementary matrices

View this lecture on YouTube

The row reduction algorithm of Gaussian elimination can be implemented by multiplying elementary matrices. Here, we show how to construct these elementary matrices, which differ from the identity matrix by a single elementary row operation. Consider the first row reduction step for the following matrix A:

$$A = \begin{pmatrix} -3 & 2 & -1 \\ 6 & -6 & 7 \\ 3 & -4 & 4 \end{pmatrix} \to \begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 3 & -4 & 4 \end{pmatrix} = M_1 A, \quad \text{where } M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

To construct the elementary matrix M_1, the number two is placed in column-one, row-two. This matrix multiplies the first row by two and adds the result to the second row.

The next step in row elimination is

$$\begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 3 & -4 & 4 \end{pmatrix} \to \begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 0 & -2 & 3 \end{pmatrix} = M_2 M_1 A, \quad \text{where } M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Here, to construct M_2 the number one is placed in column-one, row-three, and the matrix multiplies the first row by one and adds the result to the third row.

The last step in row elimination is

$$\begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 0 & -2 & 3 \end{pmatrix} \to \begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 0 & 0 & -2 \end{pmatrix} = M_3 M_2 M_1 A, \quad \text{where } M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Here, to construct M_3 the number negative-one is placed in column-two, row-three, and this matrix multiplies the second row by negative-one and adds the result to the third row.

We have thus found that

$$M_3 M_2 M_1 A = U,$$

where U is an upper triangular matrix. This discussion will be continued in the next lecture.

Problems for Lecture 13

1. Construct the elementary matrix that multiplies the second row of a four-by-four matrix by two and adds the result to the fourth row.

Solutions to the Problems

Lecture 14

LU decomposition

View this lecture on YouTube

In the last lecture, we have found that row reduction of a matrix A can be written as

$$M_3M_2M_1A = U,$$

where U is upper triangular. Upon inverting the elementary matrices, we have

$$A = M_1^{-1}M_2^{-1}M_3^{-1}U.$$

Now, the matrix M_1 multiples the first row by two and adds it to the second row. To invert this operation, we simply need to multiply the first row by negative-two and add it to the second row, so that

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarly,

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad M_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}; \quad M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad M_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Therefore,

$$L = M_1^{-1}M_2^{-1}M_3^{-1}$$

is given by

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix},$$

which is lower triangular. Also, the non-diagonal elements of the elementary inverse matrices are simply combined to form L. Our LU decomposition of A is therefore

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

Problems for Lecture 14

1. Find the LU decomposition of

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

Solutions to the Problems

Lecture 15

Solving (LU)x = b

View this lecture on YouTube

The LU decomposition is useful when one needs to solve $Ax = b$ for many right-hand-sides. With the LU decomposition in hand, one writes

$$(LU)x = L(Ux) = b,$$

and lets $y = Ux$. Then we solve $Ly = b$ for y by forward substitution, and $Ux = y$ for x by backward substitution. It is possible to show that for large matrices, solving $(LU)x = b$ is substantially faster than solving $Ax = b$ directly.

We now illustrate the solution of $LUx = b$, with

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 0 & 0 & -2 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ -7 \\ -6 \end{pmatrix}.$$

With $y = Ux$, we first solve $Ly = b$, that is

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -7 \\ -6 \end{pmatrix}.$$

Using forward substitution

$$\begin{aligned} y_1 &= -1, \\ y_2 &= -7 + 2y_1 = -9, \\ y_3 &= -6 + y_1 - y_2 = 2. \end{aligned}$$

We then solve $Ux = y$, that is

$$\begin{pmatrix} -3 & 2 & -1 \\ 0 & -2 & 5 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -9 \\ 2 \end{pmatrix}.$$

Using back substitution,

$$\begin{aligned} x_3 &= -1, \\ x_2 &= -\frac{1}{2}(-9 - 5x_3) = 2, \\ x_1 &= -\frac{1}{3}(-1 - 2x_2 + x_3) = 2, \end{aligned}$$

and we have found

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

Problems for Lecture 15

1. Using

$$A = \begin{pmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{pmatrix} \begin{pmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{pmatrix} = LU,$$

compute the solution to $Ax = b$ with

(a) $b = \begin{pmatrix} -3 \\ 3 \\ 2 \end{pmatrix}$, (b) $b = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

Solutions to the Problems

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, $\{\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n\}$, are *linearly independent* if for any scalars c_1, c_2, ..., 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 $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_n$ 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 $\{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n\}$ is the vector space consisting of all linear combinations of $\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_n$. 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

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 $v_1, v_2, \dots, v_n$ 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 $u_1, u_2, \dots, u_n$, and then normalize each vector to obtain an orthonormal basis. First, define $u_1 = v_1$. To find the next orthogonal basis vector, define

$$u_2 = v_2 - \frac{(u_1^T v_2)u_1}{u_1^T u_1}.$$

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 u_1^T, it is easy to see that $u_1^T u_2 = 0$ so that these two vectors are orthogonal.

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

$$u_3 = v_3 - \frac{(u_1^T v_3)u_1}{u_1^T u_1} - \frac{(u_2^T v_3)u_2}{u_2^T u_2}.$$

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

$$\widehat{u}_1 = \frac{u_1}{(u_1^T u_1)^{1/2}}, \quad \text{etc.}$$

Since u_k is a linear combination of $v_1, v_2, \dots, v_k$, 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

$$\text{span}\{u_1, u_2, \dots, u_k\} = \text{span}\{v_1, v_2, \dots, v_k\}.$$

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

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

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

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 underdetermined system of linear equations $Ax = b$ 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 $Av = b$, then $x = u + v$ satisfies $Ax = A(u + v) = Au + Av = 0 + b = b$. The general solution of $Ax = b$ can therefore be written as the sum of a general vector in $\text{Null}(A)$ and a particular vector that satisfies the underdetermined 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 is determined from $x_1 = 0$ and $x_2 = -x_3/2$, and we can write

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

A particular solution for the inhomogeneous system is found by solving $x_1 = 1/4$ and $x_2 + x_3/2 = -1/4$. Here, we simply take the free variable x_3 to be zero, and we find $x_1 = 1/4$ and $x_2 = -1/4$. 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

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

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 $\mathrm{rank}(A) \leq \min(m, n)$. 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 $\operatorname{rank}(A)$. Is this matrix of full rank?

Solutions to the Problems

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 $\{s_1, s_2, \ldots, s_p\}$ be an orthonormal basis for W. In general, the orthogonal projection of v onto W is given by

$$v_{\text{proj}_W} = (v^T s_1)s_1 + (v^T s_2)s_2 + \cdots + (v^T s_p)s_p;$$

and we can write

$$v = v_{\text{proj}_W} + (v - v_{\text{proj}_W}),$$

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

We can be more concrete. Using the Gram-Schmidt process, it is possible to construct a basis for V consisting of all the orthonormal basis vectors for W together with whatever remaining orthonormal vectors are required to span V. Write this basis for the n-dimensional vector space V as $\{s_1, s_2, \ldots, s_p, t_1, t_2, \ldots, t_{n-p}\}$. Then any vector v in V can be written as

$$v = a_1 s_1 + a_2 s_2 + \cdots + a_p s_p + b_1 t_1 + b_2 t_2 + b_{n-p} t_{n-p}.$$

The orthogonal projection of v onto W is in this case is seen to be

$$v_{\text{proj}_W} = a_1 s_1 + a_2 s_2 + \cdots + a_p s_p,$$

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

We can show that the vector v_{proj_W} is the unique vector in W that is closest to v. Let w be any vector in W different than v_{proj_W}, and expand w in terms of the basis vectors for W:

$$w = c_1 s_1 + c_2 s_2 + \cdots + c_p s_p.$$

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

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

Therefore, v_{proj_W} is closer to v than any other vector in W, and this fact 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 $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$?

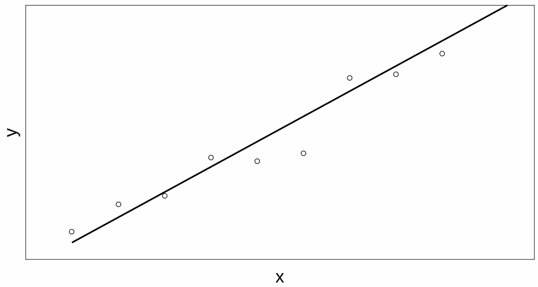
Solutions to the Problems

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 $(x_n, y_n) = (0, 1)$, $(1, 3)$, $(2, 3)$, and $(3, 4)$. If the data is to be fit by the line $y = \beta_0 + \beta_1 x$, write down the overdetermined matrix equation for β_0 and β_1.

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 $\mathrm{Ax} = \mathrm{b}$. We write $\mathrm{b} = \mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}} + (\mathrm{b} - \mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}})$, where $\mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}}$ is the projection of b onto the column space of A. Since $(\mathrm{b} - \mathrm{b}_{\mathrm{proj}_{\mathrm{Col(A)}}})$ 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 $\mathrm{Ax} = \mathrm{b}$, given by

$$\mathrm{A}^{\mathrm{T}}\mathrm{Ax} = \mathrm{A}^{\mathrm{T}}\mathrm{b}.$$

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. By manipulating the normal equations, one finds

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

Notice that the projection matrix $\mathrm{P} = \mathrm{A}(\mathrm{A}^{\mathrm{T}}\mathrm{A})^{-1}\mathrm{A}^{\mathrm{T}}$ satisfied $\mathrm{P}^2 = \mathrm{P}$, that is, two projections is the same as one.

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

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

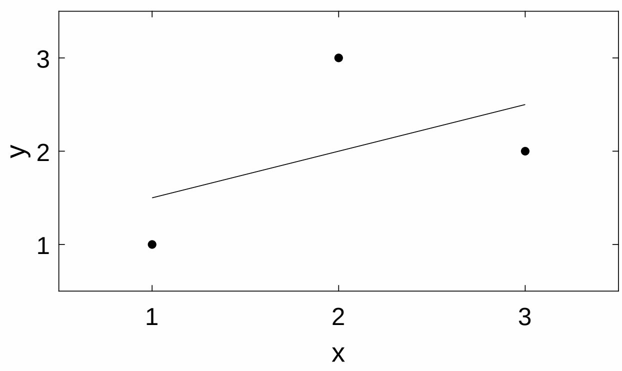
The least-squares solution is determined by solving

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

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

Week IV

Eigenvalues and eigenvectors

In this week's lectures, we will learn about determinants and the eigenvalue problem. We will learn how to compute determinants using a Laplace expansion, the Leibniz formula, or by row or column elimination. We will formulate the eigenvalue problem and learn how to find the eigenvalues and eigenvectors of a matrix. We will learn how to diagonalize a matrix using its eigenvalues and eigenvectors, and how this leads to an easy calculation of a matrix raised to a power.

Lecture 28

Two-by-two and three-by-three determinants

View this lecture on YouTube

We already showed that a two-by-two matrix A is invertible when its determinant is nonzero, where

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

If A is invertible, then the equation $\mathrm{A}x = b$ has the unique solution $x = \mathrm{A}^{-1}b$. But if A is not invertible, then $\mathrm{A}x = b$ may have no solution or an infinite number of solutions. When $\det \mathrm{A} = 0$, we say that the matrix A is singular.

It is also straightforward to define the determinant for a three-by-three matrix. We consider the system of equations $\mathrm{Ax} = 0$ and determine the condition for which $\mathrm{x} = 0$ is the only solution. With

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

one can do the messy algebra of elimination to solve for x_1, x_2, and x_3. One finds that $x_1 = x_2 = x_3 = 0$ is the only solution when $\det \mathrm{A} \neq 0$, where the definition, apart from a constant, is given by

$$\det \mathrm{A} = aei + bfg + cdh - ceg - bdi - afh.$$

An easy way to remember this result is to mentally draw the following picture:

$$\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix} - \begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}.$$

The matrix A is periodically extended two columns to the right, drawn explicitly here but usually only imagined. Then the six terms comprising the determinant are made evident, with the lines slanting down towards the right getting the plus signs and the lines slanting down towards the left getting the minus signs. Unfortunately, this mnemonic only works for three-by-three matrices.

Problems for Lecture 28

1. Find the determinant of the three-by-three identity matrix.

2. Show that the three-by-three determinant changes sign when the first two rows are interchanged.

3. Let A and B be two-by-two matrices. Prove by direct computation that $\det AB = \det A \det B$.

Solutions to the Problems

Lecture 29

Laplace expansion

View this lecture on YouTube

There is a way to write the three-by-three determinant that generalizes. It is called a Laplace expansion (also called a cofactor expansion or expansion by minors). For the three-by-three determinant, we have

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - bdi - afh$$
$$= a(ei - fh) - b(di - fg) + c(dh - eg),$$

which can be written suggestively as

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Evidently, the three-by-three determinant can be computed from lower-order two-by-two determinants, called minors. The rule here for a general n-by-n matrix is that one goes across the first row of the matrix, multiplying each element in the row by the determinant of the matrix obtained by crossing out that element's row and column, and adding the results with alternating signs.

In fact, this expansion in minors can be done across any row or down any column. When the minor is obtained by deleting the ith-row and j-th column, then the sign of the term is given by $(-1)^{i+j}$. An easy way to remember the signs is to form a checkerboard pattern, exhibited here for the three-by-three and four-by-four matrices:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}, \quad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

Example: Compute the determinant of

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

We first expand in minors down the second column:

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

We then again expand in minors down the second column, and compute the two-by-two determinant:

$$-2 \begin{vmatrix} 1 & 0 & -1 \\ 3 & 0 & 5 \\ 1 & 5 & 0 \end{vmatrix} = 10 \begin{vmatrix} 1 & -1 \\ 3 & 5 \end{vmatrix} = 80.$$

The trick here is to expand by minors across the row or column containing the most zeros.

Problems for Lecture 29

1. Compute the determinant of

$$A = \begin{pmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & 4 & 1 & 0 \\ 8 & -5 & 6 & 7 & -2 \\ -2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 3 & 2 & 0 \end{pmatrix}.$$

Solutions to the Problems

Lecture 30

Leibniz formula

View this lecture on YouTube

Another way to generalize the three-by-three determinant is called the Leibniz formula, or more descriptively, the big formula. The three-by-three determinant can be written as

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei - afh + bfg - bdi + cdh - ceg,$$

where each term in the formula contains a single element from each row and from each column. For example, to obtain the third term bfg, b comes from the first row and second column, f comes from the second row and third column, and g comes from the third row and first column. As we can choose one of three elements from the first row, then one of two elements from the second row, and only one element from the third row, there are $3! = 6$ terms in the formula, and the general n-by-n matrix without any zero entries will have $n!$ terms.

The sign of each term depends on whether the choice of columns as we go down the rows is an even or odd permutation of the columns ordered as $\{1, 2, 3, \dots, n\}$. An even permutation is when columns are interchanged an even number of times, and an odd permutation is when they are interchanged an odd number of times. Even permutations get a plus sign and odd permutations get a minus sign.

For the determinant of the three-by-three matrix, the plus terms aei, bfg, and cdh correspond to the column orderings $\{1, 2, 3\}$, $\{2, 3, 1\}$, and $\{3, 1, 2\}$, which are even permutations of $\{1, 2, 3\}$, and the minus terms afh, bdi, and ceg correspond to the column orderings $\{1, 3, 2\}$, $\{2, 1, 3\}$, and $\{3, 2, 1\}$, which are odd permutations.

Problems for Lecture 30

1. Using the Leibniz formula, compute the determinant of the following four-by-four matrix:

$$A = \begin{pmatrix} a & b & c & d \\ e & f & 0 & 0 \\ 0 & g & h & 0 \\ 0 & 0 & i & j \end{pmatrix}.$$

Solutions to the Problems

Lecture 31

Properties of a determinant

View this lecture on YouTube

The determinant is a function that maps a square matrix to a scalar. It is uniquely defined by the following three properties:

Property 1: The determinant of the identity matrix is one;

Property 2: The determinant changes sign under row interchange;

Property 3: The determinant is a linear function of the first row, holding all other rows fixed.

Using two-by-two matrices, the first two properties are illustrated by

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \qquad \text{and} \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = - \begin{vmatrix} c & d \\ a & b \end{vmatrix};$$

and the third property is illustrated by

$$\begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = k \begin{vmatrix} a & b \\ c & d \end{vmatrix} \qquad \text{and} \qquad \begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}.$$

Both the Laplace expansion and Leibniz formula for the determinant can be proved from these three properties. Other useful properties of the determinant can also be proved:

- The determinant is a linear function of any row, holding all other rows fixed;
- If a matrix has two equal rows, then the determinant is zero;
- If we add k times row-i to row-j, the determinant doesn't change;
- The determinant of a matrix with a row of zeros is zero;
- A matrix with a zero determinant is not invertible;
- The determinant of a diagonal matrix is the product of the diagonal elements;
- The determinant of an upper or lower triangular matrix is the product of the diagonal elements;
- The determinant of the product of two matrices is equal to the product of the determinants;
- The determinant of the inverse matrix is equal to the reciprical of the determinant;
- The determinant of the transpose of a matrix is equal to the determinant of the matrix.

Notably, these properties imply that Gaussian elimination, done on rows or columns or both, can be used to simplify the computation of a determinant. Row interchanges and multiplication of a row by a constant change the determinant and must be treated correctly.

Problems for Lecture 31

1. Using the defining properties of a determinant, prove that if a matrix has two equal rows, then the determinant is zero.

2. Using the defining properties of a determinant, prove that the determinant is a linear function of any row, holding all other rows fixed.

3. Using the results of the above problems, prove that if we add k times row-i to row-j, the determinant doesn't change.

4. Use Gaussian elimination to find the determinant of the following matrix:

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

Solutions to the Problems

Lecture 32

The eigenvalue problem

View this lecture on YouTube

Let A be a square matrix, x a column vector, and λ a scalar. The eigenvalue problem for A solves

$$\mathrm{Ax} = \lambda \mathrm{x}$$

for eigenvalues λ_i with corresponding eigenvectors x_i. Making use of the identity matrix I, the eigenvalue problem can be rewritten as

$$(\mathrm{A} - \lambda \mathrm{I})\mathrm{x} = 0,$$

where the matrix $(\mathrm{A} - \lambda \mathrm{I})$ is just the matrix A with λ subtracted from its diagonal. For there to be nonzero eigenvectors, the matrix $(\mathrm{A} - \lambda \mathrm{I})$ must be singular, that is,

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

This equation is called the *characteristic equation* of the matrix A. From the Leibniz formula, the characteristic equation of an n-by-n matrix is an n-th order polynomial equation in λ. For each found λ_i, a corresponding eigenvector $\mathrm{x_i}$ can be determined directly by solving $(\mathrm{A} - \lambda_i \mathrm{I})\mathrm{x} = 0$ for x.

For illustration, we compute the eigenvalues of a general two-by-two matrix. We have

$$0 = \det(\mathrm{A} - \lambda \mathrm{I}) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad-bc);$$

and this characteristic equation can be rewritten as

$$\lambda^2 - \mathrm{Tr}\,\mathrm{A}\,\lambda + \det \mathrm{A} = 0,$$

where $\mathrm{Tr}\,\mathrm{A}$ is the trace, or sum of the diagonal elements, of the matrix A.

Since the characteristic equation of a two-by-two matrix is a quadratic equation, it can have either (i) two distinct real roots; (ii) two distinct complex conjugate roots; or (iii) one degenerate real root. More generally, eigenvalues can be real or complex, and an n-by-n matrix may have less than n distinct eigenvalues.

Problems for Lecture 32

1. Using the formula for a three-by-three determinant, determine the characteristic equation for a general three-by-three matrix A. This equation should be written as a cubic equation in λ.

Solutions to the Problems

Lecture 33

Finding eigenvalues and eigenvectors (1)

View this lecture on YouTube

We compute here the two real eigenvalues and eigenvectors of a two-by-two matrix.

Example: Find the eigenvalues and eigenvectors of $\mathrm{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

The characteristic equation of A is given by

$$\lambda^2 - 1 = 0,$$

with solutions $\lambda_1 = 1$ and $\lambda_2 = -1$. The first eigenvector is found by solving $(\mathrm{A} - \lambda_1 \mathrm{I})\mathrm{x} = 0$, or

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

The equation from the second row is just a constant multiple of the equation from the first row and this will always be the case for two-by-two matrices. From the first row, say, we find $x_2 = x_1$. The second eigenvector is found by solving $(\mathrm{A} - \lambda_2 \mathrm{I})\mathrm{x} = 0$, or

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

so that $x_2 = -x_1$. The eigenvalues and eigenvectors are therefore given by

$$\lambda_1 = 1, \ \ \mathrm{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad \lambda_2 = -1, \ \ \mathrm{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The eigenvectors can be multiplied by an arbitrary nonzero constant. Notice that $\lambda_1 + \lambda_2 = \mathrm{Tr}\,\mathrm{A}$ and that $\lambda_1 \lambda_2 = \det \mathrm{A}$, and analogous relations are true for any n-by-n matrix. In particular, comparing the sum over all the eigenvalues and the matrix trace provides a simple algebra check.

Problems for Lecture 33

1. Find the eigenvalues and eigenvectors of $\begin{pmatrix} 2 & 7 \\ 7 & 2 \end{pmatrix}$.

2. Find the eigenvalues and eigenvectors of

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

Solutions to the Problems

Lecture 34

Finding eigenvalues and eigenvectors (2)

View this lecture on YouTube

We compute some more eigenvalues and eigenvectors.

Example: Find the eigenvalues and eigenvectors of $\mathrm{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

The characteristic equation of B is given by

$$\lambda^2 = 0,$$

so that there is a degenerate eigenvalue of zero. The eigenvector associated with the zero eigenvalue is found from $\mathrm{Bx} = 0$ and has zero second component. This matrix therefore has only one eigenvalue and eigenvector, given by

$$\lambda = 0, \quad \mathrm{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Example: Find the eigenvalues of $\mathrm{C} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

The characteristic equation of C is given by

$$\lambda^2 + 1 = 0,$$

which has the imaginary solutions $\lambda = \pm i$. Matrices with complex eigenvalues play an important role in the theory of linear differential equations.

Problems for Lecture 34

1. Find the eigenvalues of $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

Solutions to the Problems

Lecture 35

Matrix diagonalization

View this lecture on YouTube

For concreteness, consider a two-by-two matrix A with eigenvalues and eigenvectors given by

$$\lambda_1, \quad \mathrm{x}_1 = \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix}; \qquad \lambda_2, \quad \mathrm{x}_2 = \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix}.$$

And consider the matrix product and factorization given by

$$\mathrm{A} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} \lambda_1 x_{11} & \lambda_2 x_{12} \\ \lambda_1 x_{21} & \lambda_2 x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Generalizing, we define S to be the matrix whose columns are the eigenvectors of A, and Λ to be the diagonal matrix with eigenvalues down the diagonal. Then for any n-by-n matrix with n linearly independent eigenvectors, we have

$$\mathrm{AS} = \mathrm{S}\Lambda,$$

where S is an invertible matrix. Multiplying both sides on the right or the left by S^{-1}, we derive the relations

$$\mathrm{A} = \mathrm{S}\Lambda\mathrm{S}^{-1} \quad \text{or} \quad \Lambda = \mathrm{S}^{-1}\mathrm{AS}.$$

To remember the order of the S and S^{-1} matrices in these formulas, just remember that A should be multiplied on the right by the eigenvectors placed in the columns of S.

Problems for Lecture 35

1. Prove that the eigenvectors associated with distinct eigenvalues are linearly independent.

2. Prove that if the columns of an n-by-n matrix are linearly independent, then the matrix is invertible. (A matrix whose columns are eigenvectors corresponding to distinct eigenvalues is therefore invertible.)

Solutions to the Problems

Lecture 36

Matrix diagonalization example

View this lecture on YouTube

Example: Diagonalize the matrix $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.

The eigenvalues of A are determined from

$$\det(A - \lambda I) = \begin{vmatrix} a-\lambda & b \\ b & a-\lambda \end{vmatrix} = (a-\lambda)^2 - b^2 = 0.$$

Solving for λ, the two eigenvalues are given by $\lambda_1 = a + b$ and $\lambda_2 = a - b$. The corresponding eigenvector for λ_1 is found from $(A - \lambda_1 I)x_1 = 0$, or

$$\begin{pmatrix} -b & b \\ b & -b \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

and the corresponding eigenvector for λ_2 is found from $(A - \lambda_2 I)x_2 = 0$, or

$$\begin{pmatrix} b & b \\ b & b \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Solving for the eigenvectors and normalizing them, the eigenvalues and eigenvectors are given by

$$\lambda_1 = a + b, \quad x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad \lambda_2 = a - b, \quad x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The matrix S of eigenvectors can be seen to be orthogonal so that $S^{-1} = S^T$. We then have

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad S^{-1} = S^T = S;$$

and the diagonalization result is given by

$$\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Problems for Lecture 36

1. Diagonalize the matrix $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Solutions to the Problems

Lecture 37

Powers of a matrix

View this lecture on YouTube

Diagonalizing a matrix facilitates finding powers of that matrix. Suppose that A is diagonalizable, and consider

$$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2 S^{-1},$$

where in the two-by-two example, Λ^2 is simply

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}.$$

In general, Λ^p has the eigenvalues raised to the power of p down the diagonal, and

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

Problems for Lecture 37

1. From calculus, the exponential function is sometimes defined from the power series

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots .$$

In analogy, the matrix exponential of an n-by-n matrix A can be defined by

$$e^{\mathrm{A}} = \mathrm{I} + \mathrm{A} + \frac{1}{2!}\mathrm{A}^2 + \frac{1}{3!}\mathrm{A}^3 + \dots .$$

If A is diagonalizable, show that

$$e^{\mathrm{A}} = \mathrm{S}e^{\Lambda}\mathrm{S}^{-1},$$

where

$$e^{\Lambda} = \begin{pmatrix} e^{\lambda_1} & 0 & \dots & 0 \\ 0 & e^{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{\lambda_n} \end{pmatrix}.$$

Solutions to the Problems

Lecture 38

Powers of a matrix example

View this lecture on YouTube

Example: Determine a general formula for $\begin{pmatrix} a & b \\ b & a \end{pmatrix}^n$*, where n is a positive integer.*

We have previously determined that the matrix can be written as

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Raising the matrix to the nth power, we obtain

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}^n = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} (a+b)^n & 0 \\ 0 & (a-b)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

And multiplying the matrices, we obtain

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}^n = \frac{1}{2}\begin{pmatrix} (a+b)^n + (a-b)^n & (a+b)^n - (a-b)^n \\ (a+b)^n - (a-b)^n & (a+b)^n + (a-b)^n \end{pmatrix}.$$

Problems for Lecture 38

1. Determine $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}^n$, where n is a positive integer.

Solutions to the Problems

Appendix A

Problem solutions

Solutions to the Problems for Lecture 1

1.

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

Solutions to the Problems for Lecture 2

1. $\mathrm{B} - 2\mathrm{A} = \begin{pmatrix} 0 & -4 & 3 \\ 0 & -2 & -4 \end{pmatrix}$, $\quad 3\mathrm{C} - E:$ not defined, $\quad \mathrm{AC}:$ not defined,

$$\mathrm{CD} = \begin{pmatrix} 11 & 10 \\ 10 & 11 \end{pmatrix}, \quad \mathrm{CB} = \begin{pmatrix} 8 & -10 & -3 \\ 10 & -8 & 0 \end{pmatrix}.$$

2. $\mathrm{AB} = \mathrm{AC} = \begin{pmatrix} 4 & 7 \\ 8 & 14 \end{pmatrix}$.

3. $\mathrm{AD} = \begin{pmatrix} 2 & 3 & 4 \\ 2 & 6 & 12 \\ 2 & 9 & 16 \end{pmatrix}$, $\quad \mathrm{DA} = \begin{pmatrix} 2 & 2 & 2 \\ 3 & 6 & 9 \\ 4 & 12 & 16 \end{pmatrix}$.

Solutions to the Problems for Lecture 3

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

2. Let A be an m-by-p diagonal matrix, B a p-by-n diagonal matrix, and let $\text{C} = \text{AB}$. The ij element of C is given by

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}.$$

Since A is a diagonal matrix, the only nonzero term in the sum is $k = i$ and we have $c_{ij} = a_{ii} b_{ij}$. And since B is a diagonal matrix, the only nonzero elements of C are the diagonal elements $c_{ii} = a_{ii} b_{ii}$.

3. Let A and B be n-by-n upper triangular matrices, and let $\text{C} = \text{AB}$. The ij element of C is given by

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Since A and B are upper triangular, we have $a_{ik} = 0$ when $k < i$ and $b_{kj} = 0$ when $k > j$. Excluding the zero terms from the summation, we have

$$c_{ij} = \sum_{k=i}^{j} a_{ik} b_{kj},$$

which is equal to zero when $i > j$ proving that C is upper triangular. Furthermore, $c_{ii} = a_{ii} b_{ii}$.

Solutions to the Problems for Lecture 4

1. Let A be an m-by-p matrix, B a p-by-n matrix, and $\mathrm{C} = \mathrm{AB}$ an m-by-n matrix. We have

$$c_{ij}^{\mathrm{T}} = c_{ji} = \sum_{k=1}^{p} a_{jk} b_{ki} = \sum_{k=1}^{p} b_{ik}^{\mathrm{T}} a_{kj}^{\mathrm{T}}.$$

With $\mathrm{C}^{\mathrm{T}} = (\mathrm{AB})^{\mathrm{T}}$, we have proved that $(\mathrm{AB})^{\mathrm{T}} = \mathrm{B}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}$.

2. The square matrix $\mathrm{A} + \mathrm{A}^{\mathrm{T}}$ is symmetric, and the square matrix $\mathrm{A} - \mathrm{A}^{\mathrm{T}}$ is skew symmetric. Using these two matrices, we can write

$$\mathrm{A} = \frac{1}{2}\left(\mathrm{A} + \mathrm{A}^{\mathrm{T}}\right) + \frac{1}{2}\left(\mathrm{A} - \mathrm{A}^{\mathrm{T}}\right).$$

3. Let A be a m-by-n matrix. Then using $(\mathrm{AB})^{\mathrm{T}} = \mathrm{B}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}$ and $(\mathrm{A}^{\mathrm{T}})^{\mathrm{T}} = \mathrm{A}$, we have

$$(\mathrm{A}^{\mathrm{T}}\mathrm{A})^{\mathrm{T}} = \mathrm{A}^{\mathrm{T}}\mathrm{A}.$$

Solutions to the Problems for Lecture 5

1.

$$\mathrm{A}^{\mathrm{T}}\mathrm{A} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = \begin{pmatrix} a^2+b^2+c^2 & ad+be+cf \\ ad+be+cf & d^2+e^2+f^2 \end{pmatrix}.$$

2. Let A be a m-by-n matrix. Then

$$\mathrm{Tr}(\mathrm{A}^{\mathrm{T}}\mathrm{A}) = \sum_{j=1}^{n} (\mathrm{A}^{\mathrm{T}}\mathrm{A})_{jj} = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ji}^{\mathrm{T}} a_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2,$$

which is the sum of the squares of all the elements of A.

Solutions to the Problems for Lecture 6

1. $\begin{pmatrix} 5 & 6 \\ 4 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} 5 & -6 \\ -4 & 5 \end{pmatrix}$ and $\begin{pmatrix} 6 & 4 \\ 3 & 3 \end{pmatrix}^{-1} = \frac{1}{6}\begin{pmatrix} 3 & -4 \\ -3 & 6 \end{pmatrix}$.

2. From the definition of an inverse,

$$(AB)^{-1}(AB) = I.$$

Multiply on the right by B^{-1}, and then by A^{-1}, to obtain

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

3. We assume that A is invertible so that

$$AA^{-1} = I \quad \text{and} \quad A^{-1}A = I.$$

Taking the transpose of both sides of these two equations, using both $I^T = I$ and $(AB)^T = B^TA^T$, we obtain

$$(A^{-1})^TA^T = I \quad \text{and} \quad A^T(A^{-1})^T = I.$$

We can therefore conclude that A^T is invertible and that $(A^T)^{-1} = (A^{-1})^T$.

4. Let A be an invertible matrix, and suppose B and C are its inverse. To prove that $B = C$, we write

$$B = BI = B(AC) = (BA)C = C.$$

Solutions to the Problems for Lecture 7

1. Let Q_1 and Q_2 be orthogonal matrices. Then

$$(Q_1Q_2)^{-1} = Q_2^{-1}Q_1^{-1} = Q_2^TQ_1^T = (Q_1Q_2)^T.$$

2. Since $I\,I = I$, we have $I^{-1} = I$. And since $I^T = I$, we have $I^{-1} = I^T$ and I is an orthogonal matrix.

Solutions to the Problems for Lecture 8

1. $\mathrm{R}(-\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \mathrm{R}(\theta)^{-1}.$

2. The z-coordinate stays fixed, and the vector rotates an angle θ in the x-y plane. Therefore,

$$\mathrm{R}_z = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solutions to the Problems for Lecture 9

1.

$$P_{123} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_{132} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P_{213} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$P_{231} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_{312} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad P_{321} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

2.

$$P_{123}^{-1} = P_{123}, \quad P_{132}^{-1} = P_{132}, \quad P_{213}^{-1} = P_{213}, \quad P_{321}^{-1} = P_{321},$$
$$P_{231}^{-1} = P_{312}, \quad P_{312}^{-1} = P_{231}.$$

The matrices that are their own inverses correspond to either no permutation or a single permutation of rows (or columns), e.g., $\{1, 3, 2\}$, which permutes row (column) two and three. The matrices that are not their own inverses correspond to two permutations, e.g., $\{2, 3, 1\}$, which permutes row (column) one and three, and then two and three.

Solutions to the Problems for Lecture 10

1.

(a) Row reduction of the augmented matrix proceeds as follows:

$$\begin{pmatrix} 3 & -7 & -2 & -7 \\ -3 & 5 & 1 & 5 \\ 6 & -4 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 3 & -7 & -2 & -7 \\ 0 & -2 & -1 & -2 \\ 0 & 10 & 4 & 16 \end{pmatrix} \to \begin{pmatrix} 3 & -7 & -2 & -7 \\ 0 & -2 & -1 & -2 \\ 0 & 0 & -1 & 6 \end{pmatrix}.$$

Solution by back substitition is given by

$$\begin{aligned} x_3 &= -6, \\ x_2 &= -\frac{1}{2}(x_3 - 2) = 4, \\ x_1 &= \frac{1}{3}(7x_2 + 2x_3 - 7) = 3. \end{aligned}$$

The solution is therefore

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -6 \end{pmatrix}.$$

(b) Row reduction of the augmented matrix proceeds as follows:

$$\begin{pmatrix} 1 & -2 & 3 & 1 \\ -1 & 3 & -1 & -1 \\ 2 & -5 & 5 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Solution by back substitition is given by

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

The solution is therefore

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ -1 \end{pmatrix}.$$

Solutions to the Problems for Lecture 11

1.

(a) Row reduction proceeds as follows:

$$A = \begin{pmatrix} 3 & -7 & -2 & -7 \\ -3 & 5 & 1 & 5 \\ 6 & -4 & 0 & 2 \end{pmatrix} \to \begin{pmatrix} 3 & -7 & -2 & -7 \\ 0 & -2 & -1 & -2 \\ 0 & 10 & 4 & 16 \end{pmatrix} \to \begin{pmatrix} 3 & 0 & 3/2 & 0 \\ 0 & -2 & -1 & -2 \\ 0 & 0 & -1 & 6 \end{pmatrix} \to$$
$$\begin{pmatrix} 3 & 0 & 0 & 9 \\ 0 & -2 & 0 & -8 \\ 0 & 0 & -1 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -6 \end{pmatrix}.$$

Here, columns one, two, and three are pivot columns.

(b) Row reduction proceeds as follows:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 3 & 6 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Here, columns one and three are pivot columns.

Solutions to the Problems for Lecture 12

1.

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

Therefore,

$$\begin{pmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 2/3 & 4/3 & 1/2 \\ 1 & 2 & 1/2 \\ -3 & -5 & -1 \end{pmatrix}.$$

Solutions to the Problems for Lecture 13

1.

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix}.$$

Solutions to the Problems for Lecture 14

1.

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

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

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

Therefore,

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

Solutions to the Problems for Lecture 15

1. We know

$$A = \begin{pmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{pmatrix} \begin{pmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{pmatrix} = \text{LU}.$$

To solve $\text{LUx} = \text{b}$, we let $\text{y} = \text{Ux}$, solve $\text{Ly} = \text{b}$ for y, and then solve $\text{Ux} = \text{y}$ for x.

(a)

$$\text{b} = \begin{pmatrix} -3 \\ 3 \\ 2 \end{pmatrix}$$

The equations $\text{Ly} = \text{b}$ are given by

$$\begin{aligned} y_1 &= -3, \\ -y_1 + y_2 &= 3, \\ 2y_1 - 5y_2 + y_3 &= 2, \end{aligned}$$

with solution $y_1 = -3$, $y_2 = 0$, and $y_3 = 8$. The equations $\text{Ux} = \text{y}$ are given by

$$\begin{aligned} 3x_1 - 7x_2 - 2x_3 &= -3 \\ -2x_2 - x_3 &= 0 \\ -x_3 &= 8, \end{aligned}$$

with solution $x_3 = -8$, $x_2 = 4$, and $x_1 = 3$.

(b)

$$\text{b} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

The equations $\text{Ly} = \text{b}$ are given by

$$\begin{aligned} y_1 &= 1, \\ -y_1 + y_2 &= -1, \\ 2y_1 - 5y_2 + y_3 &= 1, \end{aligned}$$

with solution $y_1 = 1$, $y_2 = 0$, and $y_3 = -1$. The equations $\text{Ux} = \text{y}$ are given by

$$\begin{aligned} 3x_1 - 7x_2 - 2x_3 &= 1 \\ -2x_2 - x_3 &= 0 \\ -x_3 &= -1, \end{aligned}$$

with solution $x_3 = 1$, $x_2 = -1/2$, and $x_1 = -1/6$.

Solutions to the Problems for Lecture 16

1. Let v be a vector in the vector space. Then both $0v$ and $v + (-1)v$ must be vectors in the vector space and both of them are the zero vector.

2. In all of the examples, the vector spaces are closed under scalar multiplication and vector addition.

Solutions to the Problems for Lecture 17

1. Only (a) and (b) are linearly independent. (c) is linearly dependent.

Solutions to the Problems for Lecture 18

1. One possible orthonormal basis is

$$\left\{ \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ \sqrt{2} \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -\sqrt{2} \end{pmatrix} \right\}.$$

The dimension of this vector space is two.

Solutions to the Problems for Lecture 19

1.

$$u_4 = v_4 - \frac{(u_1^T v_4)u_1}{u_1^T u_1} - \frac{(u_2^T v_4)u_2}{u_2^T u_2} - \frac{(u_3^T v_4)u_3}{u_3^T u_3}.$$

Solutions to the Problems for Lecture 20

1. Define

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

Let $u_1 = v_1$. Then u_2 is found from

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

Normalizing, we obtain the orthonormal basis

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

2. Define

$$\{v_1, v_2, v_3\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Let $u_1 = v_1$. Then u_2 is found from

$$\begin{aligned} u_2 &= v_2 - \frac{(u_1^T v_2)u_1}{u_1^T u_1} \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}; \end{aligned}$$

and u_3 is found from

$$\begin{aligned} u_3 &= v_3 - \frac{(u_1^T v_3)u_1}{u_1^T u_1} - \frac{(u_2^T v_3)u_2}{u_2^T u_2} \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6} \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Normalizing the three vectors, we obtain the orthonormal basis

$$\{\widehat{\mathbf{u}}_1, \widehat{\mathbf{u}}_2, \widehat{\mathbf{u}}_3\} = \left\{ \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2\sqrt{3}}\begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Solutions to the Problems for Lecture 21

1. We bring A to reduced row echelon form:

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

The equation $Ax = 0$ with the pivot variables on the left-hand sides is given by

$$x_1 = -2x_4, \quad x_2 = x_4, \quad x_3 = x_4,$$

and a general vector in the nullspace can be written as $\begin{pmatrix} -2x_4 \\ x_4 \\ x_4 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. A basis for the null space is therefore given by the single vector $\begin{pmatrix} -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

Solutions to the Problems for Lecture 22

1. The system in matrix form is given by

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

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

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

The null space is found from the first four columns, and writing the basic variables on the left-hand side, we have the system

$$x_1 = 2x_2 + x_4, \quad x_3 = -2x_4;$$

from which we can write the general form of the null space as

$$\begin{pmatrix} 2x_2 + x_4 \\ x_2 \\ -2x_4 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

A particular solution is found from

$$x_1 - 2x_2 - x_4 = 3, \quad x_3 + 2x_4 = -2.$$

The free variables x_2 and x_4 can be set to zero, and the particular solution is determined to be $x_1 = 3$ and $x_3 = -2$. The general solution to the underdetermined system of equations is therefore given by

$$\mathrm{x} = a \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ -2 \\ 0 \end{pmatrix}.$$

Solutions to the Problems for Lecture 23

1. We find

$$A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}, \quad \text{rref}(A) = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

and $\dim(\text{Col}(A)) = 3$, with a basis for the column space given by the first three columns of A.

Solutions to the Problems for Lecture 24

1. We find the reduced row echelon from of A and A^T:

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

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

Columns one, two, and four are pivot columns of A and columns one, two, and three are pivot columns of A^T. Therefore, the column space of A is given by

$$\text{Col}(A) = \text{span}\left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \right\};$$

and the row space of A (the column space of A^T) is given by

$$\text{Row}(A) = \text{span}\left\{ \begin{pmatrix} 2 \\ 3 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

The null space of A are found from the equations

$$x_1 = -x_3 + x_5, \quad x_2 = x_3 - 2x_5, \quad x_4 = 2x_5,$$

and a vector in the null space has the general form

$$\begin{pmatrix} -x_3 + x_5 \\ x_3 - 2x_5 \\ x_3 \\ 2x_5 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 1 \\ -2 \\ 0 \\ 2 \\ 1 \end{pmatrix}.$$

Therefore, the null space of A is given by

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

The null space of A^{T} are found from the equations

$$x_1 = -2x_4, \quad x_2 = 2x_4, \quad x_3 = 5x_4,$$

and a vector in the null space has the general form

$$\begin{pmatrix} -2x_4 \\ 2x_4 \\ 5x_4 \\ x_4 \end{pmatrix} = x_4 \begin{pmatrix} -2 \\ 2 \\ 5 \\ 1 \end{pmatrix}.$$

Therefore, the left null space of A is given by

$$\mathrm{LeftNull(A)} = \mathrm{span}\left\{ \begin{pmatrix} -2 \\ 2 \\ 5 \\ 1 \end{pmatrix} \right\}.$$

It can be checked that the null space is the orthogonal complement of the row space and the left null space is the orthogonal complement of the column space. The $\mathrm{rank(A)} = 3$, and A is not of full rank.

Solutions to the Problems for Lecture 25

1. Using the Gram-Schmidt process, an orthonormal basis for W is found to be

$$\mathrm{s}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \quad \mathrm{s}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix}-2\\1\\1\end{pmatrix}.$$

The projection of v onto W is then given by

$$\mathrm{v}_{\mathrm{proj}_W} = (\mathrm{v}^{\mathrm{T}}\mathrm{s}_1)\mathrm{s}_1 + (\mathrm{v}^{\mathrm{T}}\mathrm{s}_2)\mathrm{s}_2 = \frac{1}{3}(a+b+c)\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{1}{6}(-2a+b+c)\begin{pmatrix}-2\\1\\1\end{pmatrix}.$$

When $a = 1$, $b = c = 0$, we have

$$\mathrm{v}_{\mathrm{proj}_W} = \frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{3}\begin{pmatrix}-2\\1\\1\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix};$$

and when $b = 1$, $a = c = 0$, we have

$$\mathrm{v}_{\mathrm{proj}_W} = \frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix} + \frac{1}{6}\begin{pmatrix}-2\\1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}0\\1\\1\end{pmatrix}.$$

Solutions to the Problems for Lecture 26

1.

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 3 \\ 4 \end{pmatrix}$$

Solutions to the Problems for Lecture 27

1. The normal equations are given by

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 3 \\ 4 \end{pmatrix},$$

or

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 11 \\ 21 \end{pmatrix}.$$

The solution is $\beta_0 = 7/5$ and $\beta_1 = 9/10$, and the least-squares line is given by $y = 7/5 + 9x/10$.

Solutions to the Problems for Lecture 28

1.

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \times 1 \times 1 = 1.$$

2.

$$\begin{aligned} \begin{vmatrix} d & e & f \\ a & b & c \\ g & h & i \end{vmatrix} &= dbi + ecg + fah - fbg - eai - dch \\ &= -(aei + bfg + cdh - ceg - bdi - afh) = - \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}. \end{aligned}$$

3. Let

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}.$$

Then

$$\mathrm{AB} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix},$$

and

$$\begin{aligned} \det \mathrm{AB} &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\ &= (acef + adeh + bcfg + bdgh) - (acef + adfg + bceh + bdgh) \\ &= (adeh + bcfg) - (adfg + bceh) \\ &= ad(eh - fg) - bc(eh - fg) \\ &= (ad - bc)(eh - fg) \\ &= \det \mathrm{A} \det \mathrm{B}. \end{aligned}$$

Solutions to the Problems for Lecture 29

1. We first expand in minors across the fourth row:

$$\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & 4 & 1 & 0 \\ 8 & -5 & 6 & 7 & -2 \\ -2 & 0 & 0 & 0 & 0 \\ 4 & 0 & 3 & 2 & 0 \end{vmatrix} = 2 \begin{vmatrix} 3 & 2 & 4 & 0 \\ 0 & 4 & 1 & 0 \\ -5 & 6 & 7 & -2 \\ 0 & 3 & 2 & 0 \end{vmatrix}.$$

We then expand in minors down the fourth column:

$$2 \begin{vmatrix} 3 & 2 & 4 & 0 \\ 0 & 4 & 1 & 0 \\ -5 & 6 & 7 & -2 \\ 0 & 3 & 2 & 0 \end{vmatrix} = 4 \begin{vmatrix} 3 & 2 & 4 \\ 0 & 4 & 1 \\ 0 & 3 & 2 \end{vmatrix}.$$

Finally, we expand in minors down the first column:

$$4 \begin{vmatrix} 3 & 2 & 4 \\ 0 & 4 & 1 \\ 0 & 3 & 2 \end{vmatrix} = 12 \begin{vmatrix} 4 & 1 \\ 3 & 2 \end{vmatrix} = 60.$$

Solutions to the Problems for Lecture 30

1. For each element chosen from the first row, there is only a single way to choose nonzero elements from all subsequent rows. Considering whether the columns chosen are even or odd permutations of the ordered set $\{1, 2, 3, 4\}$, we obtain

$$\begin{vmatrix} a & b & c & d \\ e & f & 0 & 0 \\ 0 & g & h & 0 \\ 0 & 0 & i & j \end{vmatrix} = afhj - behj + cegj - degi.$$

Solutions to the Problems for Lecture 31

1. Suppose the square matrix A has two zero rows. If we interchange these two rows, the determinant of A changes sign according to Property 2, even though A doesn't change. Therefore, $\det A = -\det A$, or $\det A = 0$.

2. To prove that the determinant is a linear function of row i, interchange rows 1 and row i using Property 2. Use Property 3, then interchange rows 1 and row i again.

3. Consider a general n-by-n matrix. Using the linear property of the jth row, and that a matrix with two equal rows has zero determinant, we have

$$\begin{vmatrix} \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1}+ka_{i1} & \dots & a_{jn}+ka_{in} \\ \vdots & \ddots & \vdots \end{vmatrix} = \begin{vmatrix} \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & \ddots & \vdots \end{vmatrix} + k \begin{vmatrix} \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \end{vmatrix} = \begin{vmatrix} \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & \ddots & \vdots \end{vmatrix}.$$

Therefore, the determinant doesn't change by adding k times row-i to row-j.

4.

$$\begin{vmatrix} 2 & 0 & -1 \\ 3 & 1 & 1 \\ 0 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & -1 \\ 0 & 1 & 5/2 \\ 0 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & -1 \\ 0 & 1 & 5/2 \\ 0 & 0 & 7/2 \end{vmatrix} = 2 \times 1 \times 7/2 = 7.$$

Solutions to the Problems for Lecture 32

1.

$$\begin{aligned}
0 = \det(\mathrm{A} - \lambda\mathrm{I}) &= \begin{vmatrix} a-\lambda & b & c \\ d & e-\lambda & f \\ g & h & i-\lambda \end{vmatrix} \\
&= (a-\lambda)(e-\lambda)(i-\lambda) + bfg + cdh - c(e-\lambda)g - bd(i-\lambda) - (a-\lambda)fh \\
&= -\lambda^3 + (a+e+i)\lambda^2 - (ae+ai+ei-bd-cg-fh)\lambda + aei + bfg + cdh - ceg - bdi - afh.
\end{aligned}$$

Solutions to the Problems for Lecture 33

1. Let $A = \begin{pmatrix} 2 & 7 \\ 7 & 2 \end{pmatrix}$. The eigenvalues of A are found from

$$0 = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 7 \\ 7 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 49.$$

Therefore, $2 - \lambda = \pm 7$, and the eigenvalues are $\lambda_1 = -5$, $\lambda_2 = 9$. The eigenvector for $\lambda_1 = -5$ is found from

$$\begin{pmatrix} 7 & 7 \\ 7 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

or $x_1 + x_2 = 0$. The eigenvector for $\lambda_2 = 9$ is found from

$$\begin{pmatrix} -7 & 7 \\ 7 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

or $x_1 - x_2 = 0$. The eigenvalues and corresponding eigenvectors are therefore given by

$$\lambda_1 = -5, \quad \mathrm{x}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \quad \lambda_2 = 9, \quad \mathrm{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

2. The eigenvalues are found from

$$0 = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{vmatrix} = (2-\lambda)\left((2-\lambda)^2 - 2\right).$$

Therefore, $\lambda_1 = 2$, $\lambda_2 = 2 - \sqrt{2}$, and $\lambda_3 = 2 + \sqrt{2}$. The eigenvector for $\lambda_1 = 2$ are found from

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

or $x_2 = 0$ and $x_1 + x_3 = 0$, or $\mathrm{x}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. The eigenvector for $\lambda_2 = 2 - \sqrt{2}$ is found from

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

Gaussian elimination gives us

$$\operatorname{rref}\begin{pmatrix} \sqrt{2} & 1 & 0 \\ 1 & \sqrt{2} & 1 \\ 0 & 1 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, $x_1 = x_3$ and $x_2 = -\sqrt{2}x_3$ and an eigenvector is $\mathrm{x}_2 = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}$. Similarly, the third eigenvector is $\mathrm{x}_3 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$.

Solutions to the Problems for Lecture 34

34. Let $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. The eigenvalues of A are found from

$$0 = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 1.$$

Therefore, $1 - \lambda = \pm i$, and the eigenvalues are $\lambda_1 = 1 - i$, $\lambda_2 = 1 + i$.

Solutions to the Problems for Lecture 35

1. Let λ_1 and λ_2 be distinct eigenvalues of A, with corresponding eigenvectors x_1 and x_2. Write

$$c_1\mathrm{x}_1 + c_2\mathrm{x}_2 = 0.$$

To prove that x_1 and x_2 are linearly independent, we need to show that $c_1 = c_2 = 0$. Multiply the above equation on the left by A and use $\mathrm{A}\mathrm{x}_1 = \lambda_1\mathrm{x}_1$ and $\mathrm{A}\mathrm{x}_2 = \lambda_2\mathrm{x}_2$ to obtain

$$c_1\lambda_1\mathrm{x}_1 + c_2\lambda_2\mathrm{x}_2 = 0.$$

By eliminating x_1 or by eliminating x_2, we obtain

$$(\lambda_1 - \lambda_2)c_1\mathrm{x}_1 = 0, \quad (\lambda_2 - \lambda_1)c_2\mathrm{x}_2 = 0,$$

from which we conclude that if $\lambda_1 \neq \lambda_2$, then $c_1 = c_2 = 0$ and x_1 and x_2 are linearly independent.

2. Let A be an n-by-n matrix. We have

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

Since the columns of A are linearly independent, we have $\dim(\mathrm{Col}(\mathrm{A})) = n$ and $\dim(\mathrm{Null}(\mathrm{A})) = 0$. If the only solution to $\mathrm{A}\mathrm{x} = 0$ is the zero vector, then $\det \mathrm{A} \neq 0$ and A is invertible.

Solutions to the Problems for Lecture 36

1. The eigenvalues and eigenvectors of $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ are

$$\lambda_1 = 2, \; x_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad \lambda_2 = 2 - \sqrt{2}, \; x_1 = \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}; \quad \lambda_2 = 2 + \sqrt{2}, \; x_1 = \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

Notice that the three eigenvectors are mutually orthogonal. This will happen when the matrix is symmetric. If we normalize the eigenvectors, the matrix with eigenvectors as columns will be an orthogonal matrix. Normalizing the orthogonal eigenvectors (so that $S^{-1} = S^T$) , we have

$$S = \begin{pmatrix} 1/\sqrt{2} & 1/2 & 1/2 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & 1/2 \end{pmatrix}.$$

We therefore find

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2-\sqrt{2} & 0 \\ 0 & 0 & 2+\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/2 & 1/2 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/2 & 1/2 \end{pmatrix}$$

Solutions to the Problems for Lecture 37

1.

$$
\begin{aligned}
e^{A} &= e^{S\Lambda S^{-1}} \\
&= I + S\Lambda S^{-1} + \frac{S\Lambda^2 S^{-1}}{2!} + \frac{S\Lambda^3 S^{-1}}{3!} + \dots \\
&= S\left(I + \Lambda + \frac{\Lambda^2}{2!} + \frac{\Lambda^3}{3!} + \dots\right)S^{-1} \\
&= Se^{\Lambda}S^{-1}.
\end{aligned}
$$

Because Λ is a diagonal matrix, the powers of Λ are also diagonal matrices with the diagonal elements raised to the specified power. Each diagonal element of e^{Λ} contains a power series of the form

$$
1 + \lambda_i + \frac{\lambda_i^2}{2!} + \frac{\lambda_i^3}{3!} + \dots,
$$

which is the power series for e^{λ_i}.

Solutions to the Problems for Lecture 38

1. We use the result

$$\begin{pmatrix} a & b \\ b & a \end{pmatrix}^n = \frac{1}{2}\begin{pmatrix} (a+b)^n + (a-b)^n & (a+b)^n - (a-b)^n \\ (a+b)^n - (a-b)^n & (a+b)^n + (a-b)^n \end{pmatrix}$$

to find

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}^n = \begin{pmatrix} 2^{n-1} & -2^{n-1} \\ -2^{n-1} & 2^{n-1} \end{pmatrix}.$$