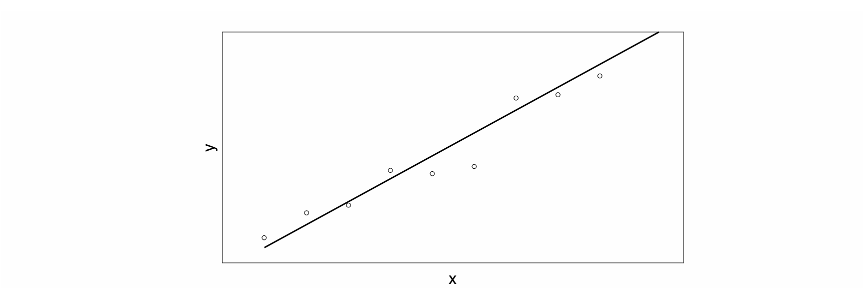
Lecture 26

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), ..., (x_n, y_n). Here, we assume that the x values are exact, and the y values are noisy. We further assume that the best fit line to the data takes the form $y = \beta_0 + \beta_1 x$. Although we know that the line will not go through all of the data points, we can still write down the equations as if it does. We have

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

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

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

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

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

Problems for Lecture 26

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

Solutions to the Problems

Lecture 27

Solution of the least-squares problem

View this lecture on YouTube

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

$$A^T Ax = A^T b.$$

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

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

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

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

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

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

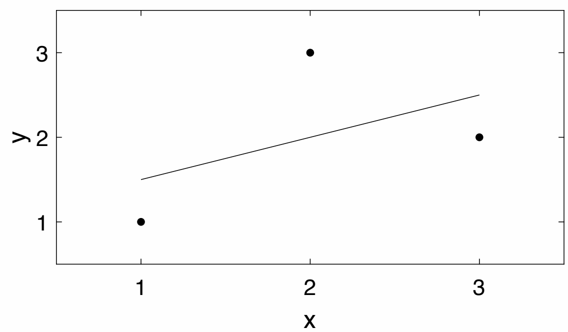
The least-squares solution is determined by solving

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

or

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

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



Solution of a toy least-squares problem.

Problems for Lecture 27

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

Solutions to the Problems

Practice quiz: Orthogonal projections

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Solutions to the Practice quiz

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

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

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. The only nonzero contribution comes from the two in the third row, and we cross out the second column and third row (and multiply by a minus sign) to obtain a three-by-three determinant:

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

We then again expand in minors down the second column. The only nonzero contribution comes from the five in the third row, and we cross out the second column and third row (and mutiply by a minus sign) to obtain a two-by-two determinant, which we then compute:

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

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:

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

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

Practice quiz: Determinants

1. The determinant of $\begin{pmatrix} -3 & 0 & -2 & 0 & 0 \\ 2 & -2 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 3 & 0 & -3 & 2 & -3 \\ -3 & 3 & 3 & 0 & -2 \end{pmatrix}$ is equal to

a) 48

b) 42

c) -42

d) -48

2. The determinant of $\begin{pmatrix} a & e & 0 & 0 \\ b & f & g & 0 \\ c & 0 & h & i \\ d & 0 & 0 & j \end{pmatrix}$ is equal to

a) $afhj + behj - cegj - degi$

b) $afhj - behj + cegj - degi$

c) $agij - beij + cefj - defh$

d) $agij + beij - cefj - defh$

3. Assume A and B are invertible n-by-n matrices. Which of the following identities is false?

a) $\det \mathrm{A}^{-1} = 1/\det \mathrm{A}$

b) $\det \mathrm{A}^{\mathrm{T}} = \det \mathrm{A}$

c) $\det(\mathrm{A}+\mathrm{B}) = \det \mathrm{A} + \det \mathrm{B}$

d) $\det(\mathrm{AB}) = \det \mathrm{A} \det \mathrm{B}$

Solutions to the Practice quiz

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

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

The eigenvectors can be multiplied by an arbitrary nonzero constant. Notice that $\lambda_1 + \lambda_2 = \operatorname{Tr} \mathrm{A}$ 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

Practice quiz: The eigenvalue problem

1. Which of the following are the eigenvalues of $\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$?

a) $\frac{3}{2} \pm \frac{\sqrt{3}}{2}$

b) $\frac{3}{2} \pm \frac{\sqrt{5}}{2}$

c) $\frac{1}{2} \pm \frac{\sqrt{3}}{2}$

d) $\frac{1}{2} \pm \frac{\sqrt{5}}{2}$

2. Which of the following are the eigenvalues of $\begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$?

a) $1 \pm 3i$

b) $1 \pm \sqrt{3}$

c) $3\sqrt{3} \pm 1$

d) $3 \pm i$

3. Which of the following is an eigenvector of $\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$?

a) $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

b) $\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}$

c) $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

d) $\begin{pmatrix} \sqrt{2} \\ 1 \\ \sqrt{2} \end{pmatrix}$

Solutions to the Practice quiz

Lecture 35

Matrix diagonalization

View this lecture on YouTube

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

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

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

Practice quiz: Matrix diagonalization

1. Let λ_1 and λ_2 be distinct eigenvalues of a two-by-two matrix A. Which of the following cannot be the associated eigenvectors?

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

b) $x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

c) $x_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $x_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

d) $x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

2. Which matrix is equal to $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{100}$?

a) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

b) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

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

d) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

3. Which matrix is equal to e^{I}, where I is the two-by-two identity matrix?

a) $\begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}$

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

c) $\begin{pmatrix} 0 & e \\ e & 0 \end{pmatrix}$

d) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

Solutions to the Practice quiz