Review of Linear Algebra and Extensions

This course will start with a review of linear algebra, from the most elementary concepts which you should have seen long ago to more advanced theorems which you might not have learned about.

In our review, we will consider the algebra of vectors and matrices.

We will also review the matrix determinant, trace and eigenvalues and eigenvectors.

We will also review linear spaces and their properties.

Advanced Topics include the Cayley-Hamilton theorem and the Perron-Frobenius theorem.

Vectors

Vectors are lists of numbers. The individual numbers are referred to as elements. The following are examples of column vectors:

$$\mathrm{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \mathrm{y} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \ \mathrm{w} = \begin{bmatrix} 3 \\ 0.1 \\ -7.2 \\ 5 \end{bmatrix}, \ \mathrm{z} = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix},$$

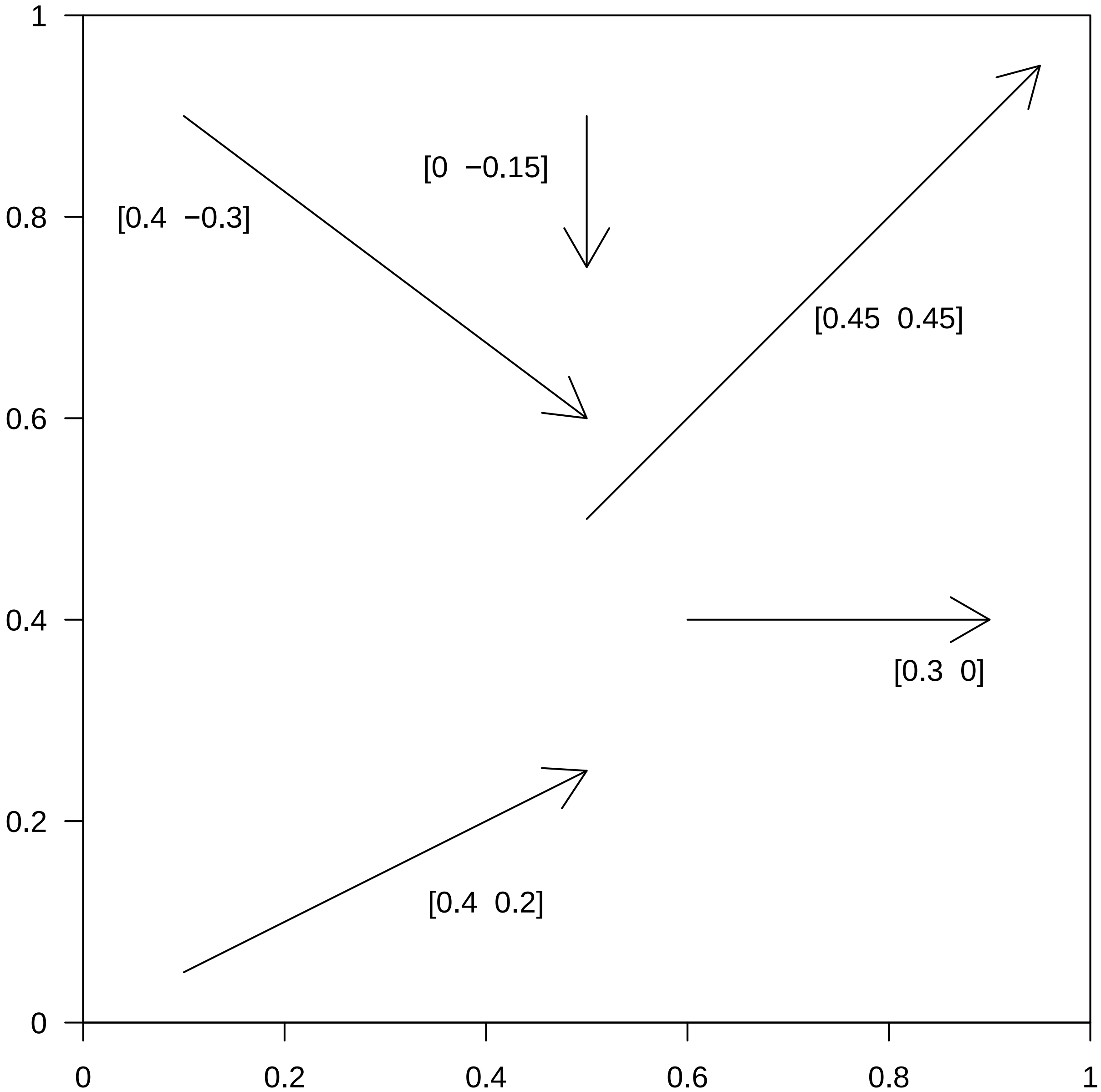
Row vectors are represented horizontally, and they are defined as the transpose of column vectors. We will use $^{\mathrm{T}}$ to denote the transpose operation.

The following are examples of row vectors:

$$\mathrm{x}^{\mathrm{T}} = [1 \ \ 1], \ \mathrm{y}^{\mathrm{T}} = [-1 \ \ 2 \ \ 0]$$

The length of a vector is its number of elements. For example, z has length 4, and y^{T} has length 3.

Graphical Display of Vectors in 2 Dimensions:



Unit Vectors

Unit vectors are vectors whose squared elements sum to 1. For example,

$$\mathbf{x} = \begin{bmatrix} 1/2 \\ \sqrt{3}/2 \end{bmatrix}$$

is a unit vector because

$$\frac{1}{2^2} + \frac{3}{2^2} = 1.$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is not a unit vector because $1^2 + 1^2 \neq 1$.

Vector Algebra – Addition

Vectors having the same length can be added together, elementwise.

For example, if

$$\mathrm{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \textbf{ and } \mathrm{y} = \begin{bmatrix} 5 \\ 9 \end{bmatrix},$$

then

$$\mathrm{x} + \mathrm{y} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}.$$

Scalar multiplication

A vector can be multiplied by a single number. The result is a vector of the same length whose elements are the products of the original elements with that single number.

For example, if

$$\mathrm{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \textbf{ and } \mathrm{y} = \begin{bmatrix} 5 \\ 9 \\ -4 \end{bmatrix},$$

then

$$7\mathrm{x} = \begin{bmatrix} 7 \\ 21 \end{bmatrix} \textbf{ and } -2\mathrm{y} = \begin{bmatrix} -10 \\ -18 \\ 8 \end{bmatrix}.$$

Inner Product

The inner product (or dot product) is the usual way of multiplying vectors which have the same length.

It is defined as the sum of the products of the corresponding elements.

The result is a single number (a scalar).

The inner product of two vectors x and y is usually denoted by the product of the transpose of x with y:

$$\mathrm{x}^{\mathrm{T}}\mathrm{y} = \sum_{i=1}^{n} x_i y_i.$$

Here, x_i denotes the ith element of the vector x which is assumed to have length n. Another notation for the inner product is

$$\langle \mathrm{x}, \mathrm{y} \rangle.$$

Inner Product – Example

If

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \text{ and } \mathbf{y} = \begin{bmatrix} 5 \\ 9 \end{bmatrix},$$

then

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = 5 + 27 = 32.$$

Note that

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{x}.$$

Orthogonal Vectors

Vectors are said to be orthogonal if their inner product is 0.

In two dimensions, this occurs if the two vectors are perpendicular to each other.

For example, if

$$\mathrm{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \text{ and } \mathrm{y} = \begin{bmatrix} -1 \\ 0.5 \end{bmatrix},$$

then

$$\mathrm{x}^{\mathrm{T}}\mathrm{y} = -1 + 1 = 0.$$

x **and** y **are orthogonal.**

Vector Norm

The norm of a vector is the square root of the inner product of the vector with itself:

$$||\mathbf{x}|| = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

For example, if

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

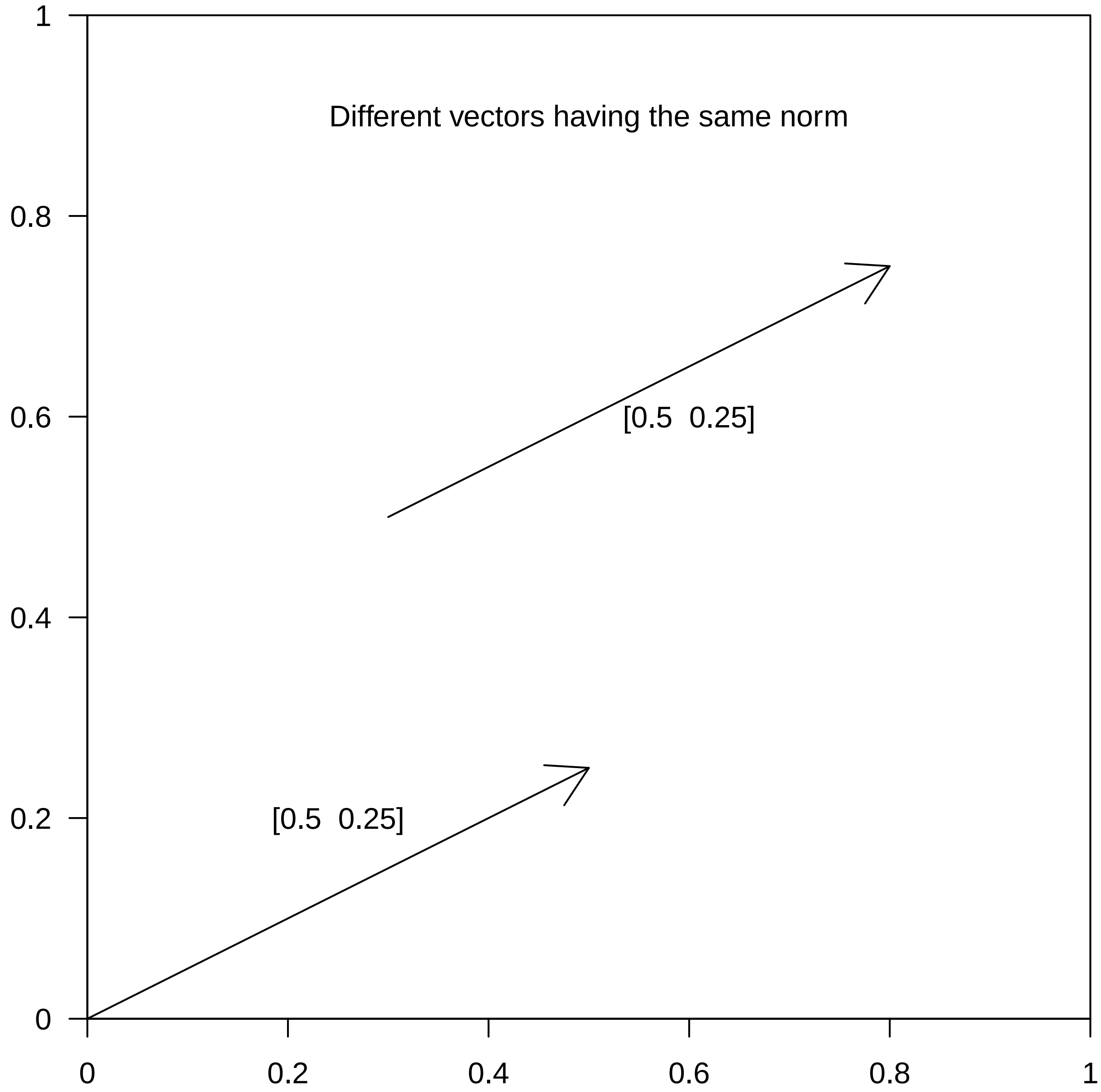
then

$$\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1 + 4 = 5.$$

$$||\mathbf{x}|| = \sqrt{5}.$$

Note that unit vectors have a norm of 1.

The Norm is the Distance Between Vector Endpoints



Projection

The projection of a vector $\mathbf{x}$ onto a vector $\mathbf{y}$ is defined as

$$\mathrm{proj}_{\mathbf{y}}\left(\mathbf{x}\right) = \frac{\mathbf{y}^{\mathbf{T}}\mathbf{x}\mathbf{y}}{||\mathbf{y}||^2}.$$

Note that if $\mathbf{x}$ and $\mathbf{y}$ are orthogonal,

$$\mathrm{proj}_{\mathbf{y}}\left(\mathbf{x}\right) = \mathbf{0}.$$

If $\mathbf{x} = \mathbf{y}$, then

$$\mathrm{proj}_{\mathbf{y}}\left(\mathbf{x}\right) = \mathbf{y}.$$

Matrices

Matrices are rectangular arrays of numbers.

A matrix with m rows and n columns is said to be an $m \times n$ matrix.

The numbers m and n are the dimensions of the matrix. For example,

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

is a 3×2 matrix, since it consists of 3 rows and 2 columns. The dimensions of A are 3 and 2.

The (i, j) element of a matrix B is denoted by $B_{i,j}$ and is located in the ith row and jth column of B. For example, $A_{2,1} = 5$.

Diagonal Matrices

A matrix $\mathbf{A}$ is called a diagonal matrix if the only nonzero elements of $\mathbf{A}$ are of the form $A_{i,i}$. These entries are the diagonal entries of $\mathbf{A}$.

For example,

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

is a diagonal matrix.

Matrix Transpose

The matrix B is said to be the transpose of a matrix $\mathbf{A}$ if $B_{j,i} = A_{i,j}$ for all i and j.

For example, if

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

then

$$\mathbf{B} = \mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 3 & 5 & 1 \\ -4 & 0 & 2 \end{bmatrix}$$

is the transpose of A.

Symmetry

A matrix $\mathbf{A}$ is said to be symmetric if $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$.

If $\mathbf{A}$ is symmetric, then $A_{i,j} = A_{j,i}$ for all i and j.

For example,

$$\mathbf{A} = \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix}$$

is a symmetric matrix.

Scalar Multiplication

If A is a matrix and b is a number, then a matrix C can be defined as

$$\mathrm{C} = b\mathrm{A}$$

where the (i, j) element of C is $C_{i,j} = bA_{i,j}$. In other words, each element of C is the product of the corresponding element of A multiplied by b.

For example, if

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

and $b = -1$, then

$$\mathrm{C} = b\mathrm{A} = -\mathrm{A} = \begin{bmatrix} -3 & 4 \\ -5 & 0 \\ -1 & -2 \end{bmatrix}$$

Matrix Addition

If A and B are $m \times n$ matrices, then their sum can be defined as

$$\mathrm{C} = \mathrm{A} + \mathrm{B}$$

where the (i, j) element of C is $C_{i,j} = A_{i,j} + B_{i,j}$. This is element-wise addition.

For example, if

$$\mathrm{A} = \begin{bmatrix} 3 & -4 \\ 5 & 0 \\ 1 & 2 \end{bmatrix} \text{ and } \mathrm{B} = \begin{bmatrix} 2 & 3 \\ 1 & 8 \\ -1 & -2 \end{bmatrix},$$

then

$$\mathrm{A} + \mathrm{B} = \begin{bmatrix} 5 & -1 \\ 6 & 8 \\ 0 & 0 \end{bmatrix}.$$

More Addition Examples

If P is the 3×2 matrix whose elements are all 0's, then we say that $\mathrm{P} = 0$ and

$$\mathrm{A} + 0 = \mathrm{A}$$

Also,

$$\mathrm{A} + (-\mathrm{A}) = 0.$$

The matrix $-\mathrm{A}$ is the additive inverse of A. Note that there is sometimes also a multiplicative inverse which will be defined later. The additive inverse always exists, but the multiplicative inverse does not always exist.

A matrix that has a multiplicative inverse is said to be invertible. A matrix that has no multiplicative inverse is said to be singular.

Matrix Multiplication

There are several ways that one could define matrix multiplication.

Here, we will use the most important definition and the one that is almost always assumed when one talks about "multiplying matrices".

Given an $\ell \times m$ matrix A and an $m \times n$ matrix B, then product AB is an $\ell \times n$ matrix C whose (i, j) element is the inner product of the ith row of A with the jth column of B:

$$C_{i,j} = \sum_{k=1}^{m} A_{i,k} B_{k,j}.$$

Note that matrix multiplication cannot be defined in this way if the dimensions of A and B do not conform. That is, the number of columns of A must match the number of rows of B in order for AB to be defined.

Example

$$\mathrm{A} = \begin{bmatrix} 3 & -4 \\ 5 & 0 \\ 1 & 2 \end{bmatrix} \text{ and } \mathrm{B} = \mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 3 & 5 & 1 \\ -4 & 0 & 2 \end{bmatrix}$$

so

$$\mathrm{AB} = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 25 & 5 \\ -5 & 5 & 5 \end{bmatrix}.$$

In this case, $\mathrm{B} = \mathrm{A}^{\mathrm{T}}$, so we have an example of the fact that AA^{T} is always a symmetric matrix.

Matrix Identity

A diagonal matrix whose diagonal entries are all 1 is called an identity matrix. It is usually denoted by the symbol I. Sometimes, we will use the symbol I_n to denote the $n \times n$ identity matrix.

The 3×3 identity matrix is given by

$$\mathbf{I}_3 = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The identity matrix has the property that multiplication with a matrix A always returns the matrix A:

$$\mathbf{IA} = \mathbf{A} = \mathbf{AI}.$$

Matrix Inverse

A matrix A is said to have an inverse (or to be invertible) if

$$\mathrm{AB} = \mathrm{BA} = \mathrm{I}$$

for some matrix B.

B is called the inverse of A and

$$\mathrm{B} = \mathrm{A}^{-1}.$$

Check that if

$$\mathrm{A} = \begin{bmatrix} 3 & -4 \\ 1 & 2 \end{bmatrix},$$

then

$$\mathrm{A}^{-1} = \begin{bmatrix} 0.2 & 0.4 \\ -0.1 & 0.3 \end{bmatrix}.$$

Relations involving Transposes and Inverses

The inverse of P^{T} is given by

$$(\mathrm{P}^{\mathrm{T}})^{-1} = (\mathrm{P}^{-1})^{\mathrm{T}}$$

provided that P is invertible. Note that if P is not invertible, then neither is P^{T}.

The transpose of $(\mathrm{P}_1\mathrm{P}_2)$ is

$$(\mathrm{P}_1\mathrm{P}_2)^{\mathrm{T}} = \mathrm{P}_2^{\mathrm{T}}\mathrm{P}_1^{\mathrm{T}}.$$

The inverse of $(\mathrm{P}_1\mathrm{P}_2)$ is

$$(\mathrm{P}_1\mathrm{P}_2)^{-1} = \mathrm{P}_2^{-1}\mathrm{P}_1^{-1}.$$

Note that in order for the inverse of $\mathrm{P}_1\mathrm{P}_2$ to exist, P_1 and P_2 must both be invertible.

Orthogonal Matrices

A matrix A is said to be orthogonal if $\mathrm{A}^T = \mathrm{A}^{-1}$.

The columns of an orthogonal matrix A must all be unit vectors and must be mutually orthogonal.

Associativity and Commutativity

Matrices are associative and commutative under the addition operation:

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$$

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$$

Matrices are associative under the multiplication operation, but not commutative in general

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$$

$$\mathbf{AB} \neq \mathbf{BA}$$

(Associativity only holds for matrices of finite dimensions; matrices with infinite dimension do not have the associativity property.)

Idempotent Matrices

A matrix A is said to be idempotent if $\mathrm{A}^2 = \mathrm{AA} = \mathrm{A}$.

For example, check that P is idempotent if

$$\mathrm{P} = \frac{1}{14}\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix},$$

then P is an idempotent matrix. That is, $\mathrm{P}^2 = \mathrm{P}$.

Solving Linear Equations

A very important application of linear algebra is in solving systems of linear equations. For example, we can represent the equations

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

as

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

where

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 1 \\ 0 & -2 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

and solve them as

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} = \begin{bmatrix} -0.5 & 1 & -0.5 \\ 0.25 & 0 & -0.25 \\ 0.5 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.5 \\ 2.0 \end{bmatrix}$$

Gaussian elimination

The way to solve linear equations which is most commonly described in linear algebra courses is Gaussian elimination. Details on this method are described, for example, at

`http://en.wikipedia.org/wiki/Gaussian_elimination`

Later in the course, we will consider better alternatives to this approach.

Trace

The trace of a matrix is the sum of its diagonal entries. It has a number of important properties, including the linearity properties:

$$\text{tr}(\mathbf{P}_1 + \mathbf{P}_2) = \text{tr}(\mathbf{P}_1) + \text{tr}(\mathbf{P}_2)$$

and

$$\text{tr}(\alpha \mathbf{P}) = \alpha \text{tr}(\mathbf{P}).$$

It also satisfies a commutativity property, even for non-commutative matrices:

$$\text{tr}(\mathbf{P}_1 \mathbf{P}_2) = \text{tr}(\mathbf{P}_2 \mathbf{P}_1)$$

To verify the last property, note that

$$\text{tr}(\mathbf{P}_1 \mathbf{P}_2) = \sum_{i=1}^{n} (\mathbf{P}_1 \mathbf{P}_2)_{ii} = \sum_{i=1}^{n} \sum_{j=1} (\mathbf{P}_1)_{ij} (\mathbf{P}_2)_{ji}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} (\mathbf{P}_2)_{ji} (\mathbf{P}_1)_{ij} = \text{tr}(\mathbf{P}_2 \mathbf{P}_1).]$$

Determinant

According to Wikipedia, "the determinant 'determines' whether the system has a unique solution (which occurs precisely if the determinant is non-zero). In this sense, determinants were first used in the Chinese mathematics textbook The Nine Chapters on the Mathematical Art:"

九章算術

The determinant of an $n \times n$ matrix $\mathbf{A}$ is given by

$$\mathsf{det}(\mathbf{A}) = \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} A_{i,\sigma_i}.$$

Matrices

Determinant

The sum is computed over all permutations σ of the set {1, 2, ..., n}. A permutation is a function that reorders this set of integers.

In any of the $n!$ summands, the term

$$\prod_{i=1}^{n} A_{i,\sigma_i}$$

is notation for the product of the entries at positions (i, σ_i), where i ranges from 1 to n:

$$A_{1,\sigma_1} \cdot A_{2,\sigma_2} \cdots A_{n,\sigma_n}.$$

Some Determinant Properties

$\det(\mathbf{A}) \neq 0$ **if and only if** $\mathbf{A}^{-1}$ **exists.**

$$\det(\mathbf{P}_1\mathbf{P}_2) = \det(\mathbf{P}_1)\det(\mathbf{P}_2).$$

$$\det(\alpha\mathbf{P}) = \alpha^n \det(\mathbf{P}).$$

$$\det(\mathbf{A}^{\mathbf{T}}) = \det(\mathbf{A}).$$

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}).$$

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}.$$

$$\det(\exp(\mathbf{A})) = \exp(\mathrm{tr}(\mathbf{A}))$$

$$\mathrm{tr}(\mathbf{A}) = \log(\det(\exp(\mathbf{A}))).$$

Eigenvalues and Eigenvectors

Consider an $n \times n$ matrix A.

If $\mathrm{Ax} = \lambda\mathrm{x}$, for some $\mathrm{x} \neq 0$ and a scalar λ, then x is said to be an eigenvector of A, and λ is an eigenvalue.

A has n eigenvalues (some may be repeated). If an eigenvalue λ is repeated k times, we say that it has algebraic multiplicity k.

Eigenvalues can be real or complex-valued.

Properties of Eigenvalues and Eigenvectors

If A is symmetric, then its eigenvalues are all real, and it will have n linearly independent eigenvectors.

The sum of the eigenvalues is equal to $\mathtt{tr}(\mathrm{A})$.

The product of the eigenvalues is equal to $\mathtt{det}(\mathrm{A})$. Thus, A^{-1} exists iff all eigenvalues of A are nonzero.

If B is an invertible matrix, then the eigenvalues of A are identical to the eigenvalues of BAB^{-1}. BAB^{-1} is called a similarity transformation.

If C is a triangular matrix, the eigenvalues appear along the diagonal of C.

Quadratic Forms

Suppose $\mathbf{A}$ is a symmetric $n \times n$ matrix and $\mathbf{x}$ is a vector of length n. Then

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$$

is a quadratic form in $\mathbf{x}$. It can be written as a quadratic polynomial in the elements of $\mathbf{x}$.

For example, if

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix},$$

then

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = 2x_1^2 + 6x_1x_2 + 4x_2^2.$$

Positive Definite Matrices

A symmetric matrix A is said to be positive definite if its quadratic forms in x are positive when $\mathrm{x} \neq 0$.

In other words, A is positive definite if and only if

$$\mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} > 0$$

whenever $\mathrm{x} \neq 0$.

Completing the square for the previous example,

$$\mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} = 2(x_1 + 1.5x_2)^2 - .5x_2^2.$$

We can make the first term of right hand side identically 0 by taking $x_1 = -1.5x_2$ for any real x_2, say $x_2 = 1$. That means that at $\mathrm{x} = [-1.5 \quad 1]$,

$$\mathrm{x}^{\mathrm{T}}\mathrm{A}\mathrm{x} = -.5(1)^2 = -.5 < 0.$$

Therefore, the matrix A is not positive definite.

Exercises

1. By completing the square, verify that

$$\mathrm{B} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

   is positive definite.

2. Show that if λ is an eigenvalue of a positive definite matrix, then $\lambda > 0$.

3. Find the eigenvalues of the matrix A used in the example, and verify that not all of them are positive. Verify that the eigenvalues of B are all positive.

4. Set

$$\mathrm{C} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}.$$

   Is C positive definite?

Linear Spaces

Vector spaces are sets of objects called vectors on which the operations of addition and scalar multiplication are defined, and in which there is a 0 vector. For each vector x in a vector space, there is a corresponding vector called $-\mathrm{x}$ which satisfies the property:

$$\mathrm{x} + (-\mathrm{x}) = 0.$$

If a vector x is multiplied by a scalar a, then $a\mathrm{x}$ is also a vector. If x and y are vectors, then $\mathrm{x} + \mathrm{y}$ is a vector.

These last two properties are linearity properties implying that vector spaces are linear spaces.

A linear combination of vectors $\mathrm{x}_1, \ldots, \mathrm{x}_k$ is any sum of the form

$$a_1\mathrm{x}_1 + a_2\mathrm{x}_2 + \cdots + a_k\mathrm{x}_k$$

where $a_1, \ldots, a_k$ are scalars.

Subspaces

A subspace is a subset of a vector space which satisfies all of the axioms of a vector space.

To check that a subset of a vector space is a subspace, it suffices to check that if x and y are members of the subset, then so is $a\mathrm{x} + b\mathrm{y}$, for all scalars a and b.

For example, consider the set of all vectors of length 2 which are orthogonal to the vector $[1 \quad 1]$. This is a subspace, since if x and y are both orthogonal to $[1 \quad 1]$, then so is $a\mathrm{x} + b\mathrm{y}$, for any scalars a and b.

Span

A set of vectors is said to span a vector space if any vector x in the vector space can be expressed as a linear combination of the spanning set.

For example, consider the subspace of vectors orthogonal to the vector $[1 \quad 1 \quad 1]$.

The vectors $[1 \quad 0 \quad -1]^{\mathrm{T}}$ and $[2 \quad -1 \quad -1]^{\mathrm{T}}$ span this subspace.

To see this, let x be any vector in this subspace, and note that $x_1 + x_2 + x_3 = 0$.

$$x = a[1 \quad 0 \quad -1]^{\mathrm{T}} + b[2 \quad -1 \quad -1]^{\mathrm{T}}$$

where $b = -x_2$ and $a = x_2 - x_3 = x_1 + 2_x 2$. The latter equality is consistent with $x_1 + x_2 + x_3 = 0$.

Linear Independence and Dependence

A set of k linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is said to be linearly independent if the only solution to the system of equations

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots a_k\mathbf{x}_k = \mathbf{0}$$

is $a_i = 0$, for $i = 1, 2, \ldots, k$.

For example, $[1 \quad 0 \quad -1]$ and $[2 \quad -1 \quad -1]$ are linearly independent.

If the set of vectors is not linearly independent, then it is said to be linearly dependent.

Basis

A set of linear independent vectors is a basis for a vector space if it spans the space.

For example, $\{[1 \quad 0 \quad -1], [2 \quad -1 \quad -1]\}$ **is a basis for the subspace of vectors orthogonal to** $[1 \quad 1 \quad 1]$**.**

Rank of a Matrix

The rank of a matrix $\mathbf{A}$ is equal to the number of linearly independent columns of $\mathbf{A}$.

It can be shown that the rank of an $n \times n$ matrix $\mathbf{A}$ is equal to $n - k$ where k is the algebraic multiplicity of the 0 eigenvalue.

Gramm-Schmidt Orthogonalization

Given a set of k linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, the Gramm-Schmidt orthogonalization procedure can be used to obtain a set of k orthogonal vectors $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ which spans the same space as that spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

The Gramm-Schmidt procedure is as follows:

$$
\begin{aligned}
\mathbf{y}_1 &= \mathbf{x}_1, & \mathbf{z}_1 &= \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} && (1)\\
\mathbf{y}_2 &= \mathbf{x}_2 - \mathrm{proj}_{\mathbf{y}_1}(\mathbf{x}_2), & \mathbf{z}_2 &= \frac{\mathbf{y}_2}{\|\mathbf{y}_2\|} && (2)\\
\mathbf{y}_3 &= \mathbf{x}_3 - \mathrm{proj}_{\mathbf{y}_1}(\mathbf{x}_3) - \mathrm{proj}_{\mathbf{y}_2}(\mathbf{x}_3), & \mathbf{z}_3 &= \frac{\mathbf{y}_3}{\|\mathbf{y}_3\|} && (3)\\
&\vdots & &\vdots && (4)\\
\mathbf{y}_k &= \mathbf{x}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{y}_j}(\mathbf{x}_k), & \mathbf{z}_k &= \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|}. && (5)
\end{aligned}
$$

Example

Apply the Gramm-Schmidt procedure to the following set of vectors to obtain an orthogonal set of vectors

$$\left\{\mathbf{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right\}.$$

$$\mathbf{y}_1 = \mathbf{x}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \mathbf{y}_2 = \mathbf{x}_2 - \mathrm{proj}_{\mathbf{y}_1}(\mathbf{x}_2) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} -$$

$$\mathrm{proj}_{\left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)}\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} -2/5 \\ 6/5 \end{pmatrix}.$$

Example

Check that the vectors u1 and u2 are orthogonal:

$$\langle \mathbf{y}_1, \mathbf{y}_2 \rangle = \left\langle \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -2/5 \\ 6/5 \end{pmatrix} \right\rangle = -\frac{6}{5} + \frac{6}{5} = 0,$$

We can then normalize the vectors by dividing out their norms

$$\mathbf{z}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ 1 \end{pmatrix} \mathbf{z}_2 = \frac{1}{\sqrt{\frac{40}{25}}} \begin{pmatrix} -2/5 \\ 6/5 \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 \\ 3 \end{pmatrix}.$$

Exercise: Analysis of an Idempotent Matrix

Suppose P is a symmetric idempotent matrix (that is, $\mathrm{P}^{\mathrm{T}} = \mathrm{P}$ and $\mathrm{P}^2 = \mathrm{P}$).

The following slides answer these questions:

1. Find the eigenvalues of P: $\lambda_1, \lambda_2, \ldots, \lambda_n$.
2. Show that the eigenvectors of P span R^n.
3. Show that there is an $n \times n$ matrix X for which $\mathrm{P} = \mathrm{X}\mathrm{D}\mathrm{X}^{-1}$ for some diagonal matrix D.
4. Show that $\mathrm{P} = \mathrm{Z}\mathrm{D}\mathrm{Z}^{\mathrm{T}}$ for an orthogonal matrix Z and for a diagonal matrix D.
5. Show that $\mathrm{tr}(\mathrm{P}) = \sum \lambda_i =$ number of nonzero eigenvalues of P.

Eigenvalues of P

Find the eigenvalues of P: $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Suppose $\mathrm{x} \neq 0$ is an eigenvector for P. Then

$$\mathbf{Px} = \lambda\mathbf{x}$$

for some scalar λ. Premultiplying both sides by P, we have

$$\mathbf{P}^2\mathbf{x} = \lambda\mathbf{Px} = \lambda^2\mathbf{x}$$

but

$$\mathbf{P}^2\mathbf{x} = \mathbf{Px} = \lambda\mathbf{x}$$

so $\lambda = \lambda^2$ which means that λ must be 0 or 1. Thus, all n eigenvalues of P must be 0's and 1's.

The eigenvectors of P span R^n

Suppose $\mathrm{x} \in R^n$. Then $\mathrm{x} = \mathrm{Px} + (\mathrm{I} - \mathrm{P})\mathrm{x}$.

Note that $\mathrm{P}(\mathrm{Px}) = \mathrm{Px}$, so Px must be an eigenvector of P (corresponding to the eigenvalue 1).

Note also that $\mathrm{P}[(\mathrm{I} - \mathrm{P})\mathrm{x}] = 0$, so $(\mathrm{I} - \mathrm{P})\mathrm{x}$ is also an eigenvector of P (corresponding to the eigenvalue 0).

Thus any $\mathrm{x} \in R^n$ can be expressed as a sum (a special case of a linear combination) of eigenvectors of P. Hence, the eigenvectors of P must span R^n.

$\mathrm{P} = \mathrm{X}\mathrm{D}\mathrm{X}^{-1}$ for diagonal D

Show that there is an $n \times n$ matrix X for which $\mathrm{P} = \mathrm{X}\mathrm{D}\mathrm{X}^{-1}$ for diagonal D.

Let X be the matrix whose n columns are the n (linearly independent) eigenvectors of P.

Then $\mathrm{P}\mathrm{X} = \mathrm{X}\mathrm{D}$ for a diagonal matrix D whose diagonal consists only of 0's and 1's (the eigenvalues of P).

Thus, $\mathrm{P} = \mathrm{X}\mathrm{D}\mathrm{X}^{-1}$.

$\mathrm{P} = \mathrm{ZDZ}^{\mathrm{T}}$ for diagonal D

Show that there is an $n \times n$ matrix Z for which $\mathrm{P} = \mathrm{ZDZ}^{\mathrm{T}}$ for diagonal D.

Suppose there are k linearly independent eigenvectors corresponding to the eigenvalue 1. These vectors span a k dimensional subspace of R^n.

It is possible to linearly transform these vectors to form an orthogonal basis for this subspace (one way to do this is via the Gramm-Schmidt orthogonalization procedure).

The $n-k$ eigenvectors corresponding to the 0 eigenvalue can be similarly transformed.

It is an easy exercise to show that the eigenvectors corresponding to the 0 eigenvalue are orthogonal to the eigenvectors for the 1 eigenvalue.

Therefore, we can use the entire set of orthogonalized eigenvectors as columns of a matrix Z which will thus be orthogonal, i.e. $\mathrm{Z}^{-1} = Z^{\mathrm{T}}$. To finish off the argument, note that $\mathrm{PZ} = \mathrm{ZD}$, and argue as above.

The Trace of $\mathbf{P}$ Equals the Number of Nonzero Eigenvalues

Show that $\text{tr}(\mathbf{P}) = \sum \lambda_i$ = number of nonzero eigenvalues of $\mathbf{P}$

$$\text{tr}(\mathbf{P}) = \text{tr}(\mathbf{X}\mathbf{D}\mathbf{X}^{\mathbf{T}}) = \text{tr}(\mathbf{X}^{\mathbf{T}}\mathbf{X}\mathbf{D}) = \text{tr}(\mathbf{D})$$

which is the total number of nonzero eigenvalues of $\mathbf{P}$.

Cayley-Hamilton Theorem

Suppose $\mathbf{A}$ is an $n \times n$ matrix. Define

$$c(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}).$$

Then $c(\mathbf{A}) = \mathbf{0}$.
Example.

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

$$c(\lambda) = (\lambda - 1)^2 \text{ so } (\mathbf{A} - \mathbf{I})^2 = \mathbf{0}.$$

Perron-Frobenius Theorem

Suppose A is an $n \times n$ non-negative matrix with the property that for some integer k, all elements of A^k are positive.

Then there exists an eigenvalue λ such that

* $\lambda > 0$.
* **its unique right and left eigenvectors are strictly positive.**
* **λ is a simple root of the characteristic equation of A.**
* **all other eigenvalues are smaller than λ in absolute value.**

Example

Consider the matrix

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

The characteristic polynomial is $c(\lambda) = (\lambda - 1)^2 - 1 = \lambda^2 - 2\lambda$.

The largest eigenvalue is $\lambda = 2$.

The right eigenvector for $\lambda = 2$ is $[1 \quad 1]^{\mathrm{T}} > 0$.

$\lambda = 2$ is a simple root of $\lambda(\lambda - 2)$.

The other eigenvalue is 0.