Chapter 16: Linear Algebra

16.1 Matrices

Define the $n \times 1$ column vector

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

and the $n \times K$ **matrix**

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nK} \end{pmatrix} = (a_{ij})_{i,j}.$$

The **transpose** of a matrix $A = (a_{ij})$ is the matrix $A^T = (a_{ji})$. For example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Matrices can be added in an obvious way, provided they are **conformable** meaning they have the same (row and column) dimensions

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1K} + b_{1K} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \cdots & a_{nK} + b_{nK} \end{pmatrix}.$$

Multiplication is a little more complicated and requires that the number of columns of A be equal to the number of rows of B. Suppose that A is $n \times K$ and B is $K \times m$, then

$$AB = \begin{pmatrix} \sum_{j=1}^K a_{1j} b_{j1} & \cdots & \sum_{j=1}^K a_{1j} b_{jm} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^K a_{nj} b_{j1} & \cdots & \sum_{j=1}^K a_{nj} b_{jm} \end{pmatrix}.$$

The resulting matrix will have dimensions equal to the number of rows of the first matrix, and the number of columns of the second matrix. Thus, the product of an $n \times K$ matrix and a $K \times m$ matrix is an $n \times m$ matrix.

AB is an $n \times m$ matrix. If $n = K$, the matrix A is **square**. If $m = n$, then we can define both AB and BA, which are both $n \times n$ matrices. Matrix multiplication is **distributive**, meaning that $(AB)C = A(BC)$. Note however that in general matrices do not **commute** so that $AB \neq BA$ (in general, when $m \neq n$, AB and BA may not even have the same dimensions).

For vectors $x, y \in \mathbb{R}^n$, we define the **inner product**

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

Two vectors are **orthogonal** if $x^T y = 0$. The Euclidean **norm** of a vector is denoted

$$||x|| = (x^T x)^{1/2} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \in \mathbb{R}^+.$$

This measures the length of the vector. We have the triangle inequality for two vectors x, y

$$||x + y|| \leq ||x|| + ||y||.$$

This says that the length of the sum is less than or equal to the sum of the lengths.

The **identity matrix** is a special kind of square matrix

$$I_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$

and satisfies $AI = IA = A$ for all square conformable matrices. The **zero matrix** 0 consists of zeros, and satisfies $A + 0 = 0 + A = 0$ for any conformable A. A square matrix B such that $AB = BA = I$ is called the **inverse** of A; not all square matrices have an inverse, but if they do then it is unique, i.e., if B^* also satisfies $AB^* = B^* A = I$, then $B = B^*$. This is left as an exercise. If the matrix A has an inverse it is called **nonsingular** otherwise it is called **singular**.

For example, a diagonal matrix

$$D = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$$

has inverse

$$D^{-1} = \begin{pmatrix} d_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n^{-1} \end{pmatrix}.$$

Diagonal matrices are quite important. They are automatically symmetric and they have the special feature that when pre and postmultiplying an $n \times n$ matrix A by a diagonal D one gets

$$
\begin{aligned}
D^{-1}AD^{-1} &= \begin{pmatrix} d_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n^{-1} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} d_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n^{-1} \end{pmatrix} \\
&= \begin{pmatrix} \frac{a_{11}}{d_1^2} & \cdots & \frac{a_{1n}}{d_1 d_n} \\ \vdots & \ddots & \vdots \\ \frac{a_{n1}}{d_1 d_n} & \cdots & \frac{a_{nn}}{d_n^2} \end{pmatrix} \\
&= (a_{ij}/d_i d_j)_{i,j}.
\end{aligned}
$$

Example 16.1.

Derive the inverse of the reverse diagonal matrix

$$
R = \begin{pmatrix} 0 & \cdots & d_1 \\ \vdots & \ddots & \vdots \\ d_n & \cdots & 0 \end{pmatrix}.
$$

Solution: The inverse matrix R^{-1} is given by

$$
R^{-1} = \begin{pmatrix} 0 & \cdots & d_n^{-1} \\ \vdots & \ddots & \vdots \\ d_1^{-1} & \cdots & 0 \end{pmatrix}.
$$

Diagonal matrices are examples of **sparse matrices**, that is, they have many zeros relative to the non-zero elements.

We next consider the relationship between different vectors.

Definition 16.1.

A set of vectors $\{x_1, \ldots, x_K\}$ with $x_i \in \mathbb{R}^n$ is called **linearly dependent** if there exist scalars $\alpha_1, \ldots, \alpha_K$ not all zero such that

$$
\alpha_1 x_1 + \cdots + \alpha_K x_K = 0.
$$

Suppose that $\alpha_i \neq 0$, then we can write

$$
x_i = \frac{-1}{\alpha_i} \sum_{j \neq i} \alpha_j x_j
$$

so that x_i is a linear combination of the other vectors.

Example 16.2.

The two vectors

$$
x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad x_2 = \begin{pmatrix} -3 \\ -3 \end{pmatrix}
$$

are linearly dependent because

$$3x_1 + x_2 = 0.$$

Definition 16.2.

A set of vectors $\{x_1, \ldots, x_K\}$ with $x_i \in \mathbb{R}^n$ is **linearly independent** if whenever for some scalars $\alpha_1, \ldots, \alpha_K$

$$\alpha_1 x_1 + \cdots + \alpha_K x_K = 0,$$

then necessarily $\alpha_1 = 0, \ldots, \alpha_K = 0$. If $X\alpha = 0$, where X is the matrix $X = (x_1, \ldots, x_K)$, then $\alpha = 0$ where $\alpha = (\alpha_1, \ldots, \alpha_K)^T$.

Definition 16.3.

The **column rank** of an $n \times K$ matrix A is the dimension of the column space of A, which is equal to the number of linearly independent columns, while the **row rank** of A is the dimension of the row space of A. The column rank and the row rank are always equal. This number is simply called the **rank** of A. A square matrix of full rank is invertible.

The rank of an $n \times K$ matrix can be: $0, 1, \ldots,$ or $\min\{n, K\}$. A full rank square matrix has rank $n = K$, and is nonsingular. The zero matrix is of rank zero by convention, otherwise the rank has to be greater than or equal to one.

The matrix

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} = ii^T,$$

where $i = (1, \ldots, 1)^T$, is of rank one because the column vector i is technically linearly independent in the sense that $\alpha i = 0$ if and only if $\alpha = 0$. In fact, suppose that v is any $n \times 1$ column vector. Then

$$A = vv^T = \begin{pmatrix} v_1^2 & v_1 v_2 & \cdots & v_1 v_n \\ v_2 v_1 & v_2^2 & \cdots & v_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n v_1 & v_n v_2 & \cdots & v_n^2 \end{pmatrix} = (a_1, \ldots, a_n)$$

is a rank one matrix. This can be seen by the following argument. Suppose that for some α_j we have

$$\sum_{j=1}^{n} \alpha_j a_j = 0.$$

Then

$$\sum_{j=1}^{n} \alpha_j v_j \times v = 0,$$

which is possible whenever $\sum_{j=1}^{n} \alpha_j v_j = 0$. But there are many candidates for this.

Let V be an $n \times K$ full rank matrix with $K \leq n$. Then the $n \times n$ matrix VV^T is of rank K, while the $K \times K$ matrix V^TV is of rank K and hence invertible. The **Trace** of a square matrix A, denoted $\text{tr}(A)$, is defined as the sum of the diagonal elements. For conformable matrices

$$\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \text{tr}(AB) = \text{tr}(BA) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}a_{ji}$$

This is true even if $AB \neq BA$. For example, when V is $n \times K$ matrix $\text{tr}(VV^T) = \text{tr}(V^TV)$. For example for column vector x and conformable square matrix A, we have

$$x^TAx = \text{tr}(x^TAx) = \text{tr}(xx^TA)$$

The **determinant** of square matrix is defined recursively. For 2×2 matrices

$$\det(A) = \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

For a 3×3 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

we have

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}),$$

where the matrices A_{ij} are defined as the original matrix with the ith row and jth column deleted

$$A_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}; \quad A_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}; \quad A_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

In general we have

$$\det(A) = \sum_{j=1}^{n} (-1)^{j-1}\det(A_{1j}),$$

where A_{ij} is the $n-1 \times n-1$ matrix with the ith row and the jth column removed so that for example

$$A_{11} = \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

For conformable matrices

$$\text{tr}(AB) = \text{tr}(BA)$$
$$\det(AB) = \det(BA).$$

This is true even if $AB \neq BA$.

A **symmetric matrix**, necessarily square, is one in which $a_{ij} = a_{ji}$ for all $i, j = 1, \ldots, n$. For example,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = A^T.$$

In general, an $n \times n$ symmetric matrix contains $n(n+1)/2$ unique elements. Matrices can be viewed as linear transformations. Suppose that A is $n \times K$ and of full rank, and u, v are $K \times 1$ vectors. Then we have $Au \in \mathbb{R}^n$, so that $A : \mathbb{R}^K \to \mathbb{R}^n$. They are linear because for $c \in \mathbb{R}$

$$\begin{aligned} A(cu) &= cAu \\ A(u+v) &= Au + Av. \end{aligned}$$

16.1.1 Linear Spaces

Definition 16.4.

The Euclidean space $\mathbb{R}^n$ is a **vector space** (linear space) since it has addition and scalar multiplication defined. A subset L is a subspace if, for any $x, z \in L \subset \mathbb{R}^n$ and any $\alpha, \beta \in \mathbb{R}$ we have

$$\alpha x + \beta z \in L$$

For example,

$$L = \left\{ x : \sum_{j=1}^{n} x_j = 0 \right\}$$

is a proper subspace of $\mathbb{R}^n$ of dimension $n-1$. Show this. **Solution**: Let $\alpha, \beta \in \mathbb{R}$ and let $x, z \in L$. Then, because $x, z \in L$,

$$\sum_{j=1}^{n} x_j = 0 \quad \text{and} \quad \sum_{j=1}^{n} z_j = 0$$

Let $y = \alpha x + \beta z$. Then,

$$\sum_{j=1}^{n} y_j = \sum_{j=1}^{n} (\alpha x_j + \beta z_j) = \alpha \sum_{j=1}^{n} x_j + \beta \sum_{j=1}^{n} z_j = \alpha \cdot 0 + \beta \cdot 0 = 0$$

Thus, $y \in L$. This proves that L is a subspace of $\mathbb{R}^n$. We will show that the dimension of L is $n-1$ by finding a set of $n-1$ linearly independent vectors in L. For $i = 1, 2, \ldots, n-1$, let

$$v_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix},$$

where v_i is an $n \times 1$ column vector with zeros everywhere, except the i-th element is 1 and the n-th element is -1. Clearly, $\{v_i\}_{i=1}^{n-1} \in L$. We will now show that these vectors are linearly independent. Suppose that

$$\sum_{i=1}^{n-1} \alpha_i v_i = \mathbf{0}.$$

Then $\alpha_i = 0$ for $i = 1, 2, \ldots, n-1$. If we add an additional linearly independent vector to the set, for example the vector

$$v_n = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix},$$

we will have a set of n linearly independent vectors in $\mathbb{R}^n$.

Definition 16.5.

A set of vectors $x_1, \ldots, x_K \in \mathbb{R}^n$ generate a subspace of $\mathbb{R}^n$ called the **span** of $x_1, \ldots, x_K$

$$C(x_1, \ldots, x_K) = \{\alpha_1 x_1 + \cdots + \alpha_K x_K : \alpha_1, \ldots, \alpha_K \in \mathbb{R}\} = \{X\alpha, \quad \alpha \in \mathbb{R}^K\}.$$

Definition 16.6.

The **null space** of $x_1, \ldots, x_K$

$$N(x_1, \ldots, x_K) = \{\alpha_1 x_1 + \cdots + \alpha_K x_K = 0 : \alpha_1, \ldots, \alpha_K \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^n$. This is also denoted by $C^{\perp}(X)$ and called the **orthocomplement** of $C(X)$.

Definition 16.7.

A **basis** for a space L is a set of linearly independent vectors $x_1, \ldots, x_K \in \mathbb{R}^n$ such that for any $x \in L$, there exist scalars $\alpha_1, \ldots, \alpha_K$ such that

$$x = \alpha_1 x_1 + \cdots + \alpha_K x_K$$

The dimension of the space L is K.

If the vectors $x_1, \ldots, x_K$ are linearly independent then the dimension of $C(x_1, \ldots, x_K)$ is K and the dimension of $N(x_1, \ldots, x_K)$ is $n - K$. A basis is not unique.

Definition 16.8.

Let $x_1, \ldots, x_K \in \mathbb{R}^n$ be a basis for the space L. Suppose that $x_i^T x_j = 0$ for $i \neq j$ and $x_i^T x_i = 1$. Then the basis is **orthonormal** and there is only one such orthonormal basis.

Example 16.3.

For example $\mathbb{R}^n$ has orthonormal basis $e_1 = (1, 0, \ldots, 0)^T, \ldots, e_n = (0, 0, \ldots, 0, 1)^T$.

16.1.2 Eigenvectors and Eigenvalues

We next define the concept of **Eigenvectors** and **Eigenvalues**.

Definition 16.9.

For a real matrix A with dimensions $n \times n$, a vector $u \in \mathbb{R}^n$ with $u \neq 0$ and scalar $\lambda \in \mathbb{R}$ are called an **eigenvector** and **eigenvalue** respectively of the matrix A if

$$Au = \lambda u.$$

Clearly, $u = 0$ trivially satisfies this equation for any λ, which is why we don't consider it. The interpretation of eigenvector is a direction that is invariant under the transformation A. Most vectors u will have their "direction" changed by a given transformation A, the eigenvectors are the special ones that are unchanged in direction and just scaled according to λ by the transformation. For an identity matrix $A = I_n$, the eigenvectors are any vectors $u \in \mathbb{R}^n$ and the eigenvalues are all $\lambda = 1$ because $Iu = u$ for all $u \in \mathbb{R}^n$, the identity transformation does not change anything. Do eigenvalues/eigenvectors always exist, and how many of them are there?

We can write the equation $Au = \lambda u$ as

$$(A - \lambda I)u = 0,$$

where 0 is a vector of zeros. If $u \neq 0$, then $A - \lambda I$ must be singular otherwise we would have a contradiction. Therefore,

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

which gives one way of finding the eigenvalues. The left hand side of the equation is an nth order polynomial in λ, denoted $p_n(\lambda)$, called the **characteristic polynomial**. We know from the mathematical analysis of polynomial equations that $p_n(\lambda) = 0$ will always have at least one solution, although some of them may be complex valued. In general, there may be one solution, two solutions, or even as many as n distinct solutions, but no more than that, that is, we can write

$$p_n(\lambda) = (\lambda_1 - \lambda)^{k_1} \times \cdots \times (\lambda_p - \lambda)^{k_p},$$

where $\lambda_j, j = 1, \ldots, p$ are distinct solutions and k_j are called the multiplicity of the solution, where $p \leq n$ and $\sum_{j=1}^{p} k_j = n$. For example, $k_j = 1$ for all j is the case where there are n distinct solutions $\lambda_1, \ldots, \lambda_n$. The **Cayley Hamilton Theorem** says that the matrix A satisfies its own characteristic polynomial, i.e. $p_n(A) = 0$. The eigenvectors are then found by finding the vectors $u \in \mathbb{R}^n$ that solve $(A - \lambda I)u = 0$.

For real symmetric matrices A the eigenvalues are all real valued because in that case the polynomial coefficients satisfy the conditions to guarantee real solutions. A proof of this result is simple but requires complex numbers.

Example 16.4.

In the 2×2 case we may use this to find explicitly the eigenvalues and hence the eigenvectors. Suppose that

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

Then we must solve the quadratic equation

$$\det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + ad - bc = 0$$

for λ. There will always be a solution in the complex plane (fundamental theorem of algebra); the solution will be real provided

$$\Delta = (a+d)^2 - 4(ad - bc) \geq 0.$$

The general solutions are $\frac{1}{2}(a + d + \sqrt{\Delta})$ and $\frac{1}{2}(a + d - \sqrt{\Delta})$. In the special case that $c = b = 0$ it is clear that $\lambda = a$ and $\lambda = d$ are the solutions. In the special case that $a = d = 1$ and $b = c = \rho$, the eigenvalues are $1 + \rho$ and $1 - \rho$. How to find the eigenvectors? We look for solutions to

$$\begin{pmatrix} -\rho & \rho \\ \rho & -\rho \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

That is $v = u$, which implies that any multiple of $(1, 1)$ is a candidate eigenvector. For the other eigenvector we have to solve

$$\begin{pmatrix} \rho & \rho \\ \rho & \rho \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

so that $u = -v$, and any multiple of $(1, -1)$ is a candidate.

If there are n distinct eigenvalues, then eigenvectors are unique up to a scaling factor. Note that if $Au = \lambda u$, then $Au' = \lambda u'$, where $u' = ku$ for any scalar k.

Theorem 16.1.

Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof: Suppose that $Au = \lambda u$ and $Av = \mu v$. Then $v^T Au = \lambda v^T u$ and $u^T Av = \mu u^T v$. Therefore,

$$0 = v^T Au - u^T Av = \lambda v^T u - \mu u^T v = (\lambda - \mu) v^T u,$$

which means that $v^T u = 0$. □

Let λ^* be such that $p_n(\lambda^*) = 0$ and suppose that $u_1, \ldots, u_m$ are vectors such that $(A - \lambda^* I)u_i = 0$, then for any $v = \sum_{i=1}^m \alpha_i u_i$ with $\alpha_i \in \mathbb{R}$, we have

$$(A - \lambda^* I)v = 0.$$

Some matrices have repeated eigenvalues, in which case the corresponding eigenvectors form a vector space of the same dimensions as the cardinality of multiplicity.

Example 16.5.

Consider the real symmetric matrix

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

We have two distinct eigenvalues $\lambda = 1 - \rho$ and $\mu = 1 + 2\rho$. Define the vectors:

$$u_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}; \quad u_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}; \quad u_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then check that u_1 and u_2 are distinct eigenvectors associated with the eigenvalue λ, while u_3 is the eigenvector associated with μ. We have $u_1^T u_3 = u_2^T u_3 = 0$ but $u_1^T u_2 = 1 \neq 0$. Define the space

$$L = \{\alpha u_1 + \beta u_2 : \alpha, \beta \in \mathbb{R}\}.$$

This space has dimensions two since u_1, u_2 are linearly independent. Define $e_1 = u_1, e_2 = u_1 - 2u_2$, and $e_3 = u_3$. Then $\{e_1, e_2\}$ is an orthogonal basis for L. Note also that e_1, e_2 are eigenvectors associated with λ. Finally, we may scale the vectors e_j to have unit length; for example let $e_3^* = e_3/\sqrt{3}$.

Theorem 16.2.

Suppose that A is a real symmetric matrix, then it possesses an **Eigendecomposition**, whereby it can be written as

$$A = Q\Lambda Q^{-1}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

where $Q = (q_1, \ldots, q_n)$ are linearly independent eigenvectors of A associated with the eigenvalues in Λ. By convention we organize the eigenvalues in decreasing order so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Moreover, there exists a unique orthonormal matrix U

$$A = U\Lambda U^T = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

$$UU^T = (u_i^T u_j)_{i,j} = I_n = U^T U$$

where $u_i, \lambda_i, i = 1, \ldots, n$ are the real eigenvectors and real valued eigenvalues (not necessarily distinct) of the matrix A. The matrix U is orthonormal and satisfies $AU = \Lambda U$. We may equivalently write

$$U^T A U = \Lambda$$

In general finding eigenvectors requires we solve the equation $Bu = 0$ for a singular matrix $B = A - \lambda I$. There are many methods for doing this but they are quite involved to describe, so we refer the reader to elsewhere. In low dimensional cases as we have seen it is easy to see how to find the solutions. It follows from this theorem that

$$\text{tr}(A) = \text{tr}(U\Lambda U^T) = \text{tr}(U^T U\Lambda) = \sum_{i=1}^{n} \lambda_i$$

$$\det(A) = \det(U\Lambda U^T) = \det(U^T U\Lambda) = \prod_{i=1}^{n} \lambda_i.$$

It also follows that if $c \in \mathbb{R}$, the eigenvalues of $A + cI_n$ are $\lambda_1 + c, \ldots, \lambda_n + c$ and the eigenvectors of $A + cI_n$ are the same as the eigenvectors of A since for eigenvalue, eigenvector pair λ, u we have

$$(A + cI_n)u = Au + cu = (\lambda + c)u.$$

Corollary 16.1.

(Singular Value Decomposition) Suppose that A is an $n \times K$ real matrix. Then there exists a factorization, called a **singular value decomposition** of A of the form

$$A = USV^T,$$

where U is an $n \times n$ orthonormal matrix, V is a $K \times K$ orthonormal matrix, and S is an $n \times K$ matrix with non-negative real numbers on the diagonal, of the form

$$S = \begin{pmatrix} s_1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & \cdots & s_K & \cdots & 0 \end{pmatrix} = \mathrm{diag}\{s_1, \ldots, s_K\} | 0_{n-K \times K}.$$

It follows that for any matrix A we obtain $A^T A = VS^2V^T = V\Lambda V^T$, where Λ is the matrix of eigenvalues.

Definition 16.10.

A positive (semi)definite matrix (psd) A satisfies

$$x^T Ax = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \geq 0$$

for all vectors x. A strictly positive definite matrix (pd) is one for which $x^T Ax > 0$ for all x.

A negative semi-definite matrix satisfies $x^T Ax \leq 0$ for all vectors x. A matrix may be indefinite, i.e., $x^T Ax > 0$ for some x and $x^T Ax < 0$ for other x. The definiteness question can be answered for a real symmetric matrix from its eigenvalues. From the eigendecomposition we see that for any vector x

$$x^T Ax = x^T U \Lambda U^T x = y^T \Lambda y = \sum_{j=1}^n \lambda_j y_j^2,$$

where $y = U^T x$ and $x = Uy$ are in one to one correspondence. A real symmetric matrix is psd if all its eigenvalues are nonnegative.

Example 16.6.

Consider the matrix

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

Then, for any $x = (x_1, x_2)^T$ we have

$$x^T Ax = x_1^2 + x_2^2 + 2x_1 x_2 \rho.$$

The eigenvalues of A are $1 - \rho, \rho + 1$. Therefore, it is psd if and only if $\rho \in [-1, 1]$.

Definition 16.11.

A matrix $A \geq B$ if and only if $A - B$ is psd.

The matrix order is only a partial order, meaning not all matrices can be compared, as the following example illustrates.

Example 16.7.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 1/4 \end{pmatrix}$$

$$A - B = \begin{pmatrix} -1 & 0 \\ 0 & 3/4 \end{pmatrix}$$

Scalar functions of a matrix give a total order, but different scalar functions yield different rankings

$$2 = \mathrm{tr}(A) < \mathrm{tr}(B) = 9/4$$
$$1 = \det(A) > \det(B) = 1/2.$$

Theorem 16.3.

Let A be a real symmetric matrix. Then, the largest eigenvalue of A denoted λ_1 or $\lambda_{\max}(A)$, is the value of the optimized criterion in the constrained optimization problem

$$\max_{x:x^Tx=1} x^T A x$$

and the optimizing choice of x is the corresponding eigenvector.

Proof: For every eigenvalue λ there is x such that $Ax = \lambda x$, which implies that

$$x^T A x = x^T \lambda x = \lambda$$

so just take the largest such. $\square$

Likewise the smallest eigenvalue is defined through a minimization problem. That is,

$$\lambda_{\min}(A) = \min_{x:x^Tx=1} x^T A x.$$

The eigendecomposition can be used to define functions of matrices. For example, the square root of a symmetric positive definite matrix is

$$X = A^{1/2} = U\Lambda^{1/2}U^T = U \begin{pmatrix} \lambda_1^{1/2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^{1/2} \end{pmatrix} U^T.$$

We have $XX = U\Lambda^{1/2}U^TU\Lambda^{1/2}U^T = U\Lambda^{1/2}\Lambda^{1/2}U^T = U\Lambda U^T$. One may define exponentials and logarithms of matrices in this way.

The eigendecomposition is also useful in establish bounds on quadratic forms.

Theorem 16.4.

Suppose that the $n \times K$ matrix A has full rank with $K \leq n$. Then for some $C > 0$ we have for all $x \in \mathbb{R}^K$

$$C||x|| \leq ||Ax|| \leq \frac{1}{C}||x||.$$

Proof: The matrix A is not necessarily symmetric or even square, but A^TA and AA^T are symmetric. Because A is of full rank, so is $A^T A$. Therefore, we have.

$$\lambda_{min}(A^T A) = inf(x/||x||)^T A^T A(x/||x||) \geq C > 0$$

which implies that for any x,

$$x^T A^T Ax \geq C||x||^2$$

as required. Similarly for the upper bound.

16.1.3 Applications

Definition 16.12.

Suppose that $X \in \mathbb{R}^d$ is a random variable. The covariance matrix of the random vector X is defined as

$$\begin{aligned}\text{cov}(X) = E[(X-\mu)(X-\mu)^T] = E[XX^T] - \mu\mu^T &= \Sigma \\ &= \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{1d} & \cdots & \sigma_{dd} \end{pmatrix},\end{aligned}$$

where $\mu = E(X)$ and $\sigma_{ij} = \text{cov}(X_i, X_j)$.

Because Σ is a covariance matrix this means that it is symmetric (check this) and positive semi-definite, because for any vector $w \in \mathbb{R}^d$

$$0 \leq \text{var}(w^T X) = E[w^T(X-\mu)(X-\mu)^T w] = w^T \Sigma w.$$

Because it is real and symmetric we may define the inverse and the square root of Σ, provided it is strictly positive definite. In fact, the matrix $\Sigma^{-1/2}$ can be defined from the eigendecomposition, as explained above. This is the matrix equivalent of “one over the standard deviation” and it allows us to “standardize” vector random variables.

Theorem 16.5.

Let $Z = \Sigma^{-1/2}(X-\mu)$. Then $EZ = 0$ and $\text{cov}(Z) = I_d$.

Proof: We have

$$\begin{aligned}E(ZZ^T) &= E[\Sigma^{-1/2}(X-\mu)(X-\mu)^T\Sigma^{-1/2}] \\ &= \Sigma^{-1/2}E[(X-\mu)(X-\mu)^T]\Sigma^{-1/2} \\ &= \Sigma^{-1/2}\Sigma\Sigma^{-1/2} \\ &= I_d. \quad \square\end{aligned}$$

Example 16.8.

MULTIVARIATE NORMAL. We say that $X = (X_1, \ldots, X_k) \sim MVN_k(\mu, \Sigma)$, when

$$f_X(x|\mu, \Sigma) = \frac{1}{(2\pi)^{k/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right),$$

where Σ is a $k \times k$ covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix}$$

and $\det(\Sigma)$ is the determinant of Σ. The characteristic function of X is

$$E[\exp(it^T X)] = \varphi_X(t|\mu, \Sigma) = \exp\left(-t^T\mu + \frac{1}{2}t^T\Sigma t\right)$$

for any $t \in \mathbb{R}^k$.

Theorem 16.6.

Suppose that we partition $X = (X_a^T, X_b^T)^T$ and

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}; \quad \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

a. If $X \sim MVN_k(\mu, \Sigma)$ then $X_a \sim N(\mu_a, \Sigma_{aa})$. This is shown by integration of the joint density with respect to the other variables.
b. The conditional distributions of X are Gaussian too, i.e.,

$$f_{X_a|X_b}(X_a) \sim N(\mu_{X_a|X_b}, \Sigma_{X_a|X_b}),$$

where the conditional mean vector and conditional covariance matrix are given by

$$\begin{aligned} \mu_{X_a|X_b} &= E(X_a|X_b) = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(X_b - \mu_b) \\ \Sigma_{X_a|X_b} &= \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}. \end{aligned}$$

c. If Σ is diagonal, then $X_1, \ldots, X_k$ are mutually independent. In this case

$$\begin{aligned} \det(\Sigma) &= \sigma_{11} \times \cdots \times \sigma_{kk} \\ -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) &= -\frac{1}{2}\sum_{l=1}^{k}\frac{(x_l - \mu_l)^2}{\sigma_{ll}} \end{aligned}$$

so that

$$f_X(x|\mu, \sigma) = \frac{1}{\sigma_{11}^{1/2}\sqrt{2\pi}}exp(-\frac{1}{2}(\frac{x_1 - \mu_1}{\sigma_{11}})^2)\ldots\frac{1}{\sigma_{kk}^{1/2}\sqrt{2\pi}}exp(-\frac{1}{2}(\frac{x_k - \mu_k}{\sigma_{kk}})^2)$$

Example 16.9.

Suppose that $X \in \mathbb{R}^d$ and $X \sim N(\mu, \Sigma)$. We have the eigendecomposition where $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$

$$\Sigma = \sum_{i=1}^{d} \lambda_i u_i u_i^T.$$

The first **Principal Component** of the random vector X is the scalar combination $u^T X$ such that

$$\text{var}(u^T X) = E[u^T(X-\mu)(X-\mu)^T u] = u^T \Sigma u$$

is maximized subject to $u^T u = 1$, i.e., u is the eigenvector of Σ corresponding to the largest eigenvalue of Σ: u_1. The largest eigenvalue $\lambda_1 = \text{var}(u_1^T X)$. Can define

$$u_1^T X, \ldots, u_n^T X$$

as the Principal components of X in decreasing order of importance. The field of Principal Components Analysis originates with Pearson (1901) and is now very widely applied to understand multivariate statistics. See Fig. 16.1.

Example 16.10.

Suppose that $X \in \mathbb{R}^d$ and

$$X \sim N(\mu, \Sigma).$$

The parameters θ contain all the elements of $\mu = (\mu_1, \ldots, \mu_d)$ and the unique elements of the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{1d} & \cdots & \sigma_{dd} \end{pmatrix}$$

Suppose we have a sample $X_1, \ldots, X_n$. The log likelihood is

$$l(\theta|X^n) = -\frac{nd}{2}\log 2\pi - \frac{n}{2}\log\det(\Sigma) - \frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)^T\Sigma^{-1}(X_i-\mu)$$

$$\begin{aligned} l(\theta|X^n) &= -\frac{nd}{2}\log 2\pi - \frac{n}{2}\log\det(\Sigma) - \frac{1}{2}\sum_{i=1}^{n}\text{tr}((X_i-\mu)(X_i-\mu)^T\Sigma^{-1}) \\ &= -\frac{nd}{2}\log 2\pi - \frac{n}{2}\log\det(\Sigma) - \frac{n}{2}\text{tr}(S\Sigma^{-1}) \\ &- \frac{n}{2}(\bar{X}-\mu)^T\Sigma^{-1}(\bar{X}-\mu), \end{aligned}$$

where the sample mean and sample covariance matrix respectively are:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i; \quad S = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})(X_i-\bar{X})^T.$$

This follows because

$$\sum_{i=1}^{n}(X_i-\mu)(X_i-\mu)^T = \sum_{i=1}^{n}(X_i-\bar{X})(X_i-\bar{X})^T + n(\bar{X}-\mu)(\bar{X}-\mu)^T$$

It can be shown that $\bar{X}$ is the MLE of μ (this is easy) and S is the MLE of Σ (this is hard).

Example 16.11.

The motivating idea behind the search engine Google is that you want the first items returned by a search to be the most important items. One way is to count the number of sites that contain a link to a given site, and the site that is linked to the most is then the most important site. This has the drawback that all links are treated as equal. If your site is referenced from the home page of Justin Bieber, it counts no more than if it's referenced by an unknown person with no fan base. The Page rank method takes account of the importance of each website in terms of its links. Suppose there are N web sites, the page rank vector r satisfies

$$r = \frac{1-d}{N} i + dAr,$$

where i is a vector of ones, while $A = (A_{ij})$ is the **Adjacency matrix** with $A_{ij} = 0$ if page j does not link to page i, and normalized such that, for each j, $\sum_{i=1}^{n} A_{ij} = 1$. Here, $d \leq 1$ is a dampening factor. When $d = 1$ we have

$$Ar = r$$

so that r is the eigenvector of A corresponding to unit eigenvalue.

Example 16.12.

Input output analysis. Suppose that the economy has n sectors. Each sector produces x_i units of a single homogeneous good. Assume that the jth sector, in order to produce 1 unit, must use a_{ij} units from sector i. Furthermore, assume that each sector sells some of its output to other sectors (intermediate output) and some of its output to consumers (final output, or final demand). Call final demand in the ith sector d_i. Then we might write

$$x - Ax = d$$

where $A = (a_{ij})$. We can solve this equation by

$$x = (I - A)^{-1} d$$

provided the inverse exists.

Theorem 16.7.

Let A be an $n \times m$ matrix of real numbers. Either there exists $\pi \in \mathbb{R}^m$ such that $A\pi \geq 0$ for all elements, or there exists $\alpha \in \mathbb{R}^n$ such that $A^T \alpha < 0$ for all elements.

Example 16.13.

Dirty Harry offers the following odds on the Premiership title: Arsenal 2 to 1, Manchester City 3 to 1, Chelsea 4 to 1 and Manchester United 5 to 1. You may assume that there are only 4 teams with positive probability of winning. Is it possible to find a combination of bets at these odds that you will surely win no matter who wins the Premiership? We can represent the payoff matrix for the bettor as

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

This matrix is positive definite and has eigenvalues $1.1961136, 4.4922513, 5.6077247, 6.7039104$ with first eigenvector

$$u = \begin{pmatrix} 0.687 \\ 0.507 \\ 0.401 \\ 0.332 \end{pmatrix}$$

that is $Au = \lambda u > 0$. This says that you should bet in proportion to $u_i / \sum_{i=1}^{4} u_i$, i.e., 0.357 on Arsenal, 0.263 on Manchester City, 0.208 on Chelsea, and 0.172 on Manchester United. No matter what you will make 1.196.

16.2 Systems of Linear Equations and Projection

We next consider linear equations and their solution. Consider the system of equations

$$Ax = y,$$

where A is $n \times K$ and y is $n \times 1$, and both are given. We seek the $K \times 1$ solution vector x. In general, there may be no solution to these equations, many solutions, or a unique solution. We suppose that A is of full rank. There are several cases.

1. Suppose that $n = K$. Then, since A is nonsingular we may write the unique solution

   $$x = A^{-1}y$$

2. Second, suppose that $n < K$. In this case, there are multiple solutions. Suppose that $y = 0$. Then for any $K \times 1$ vector w

   $$x = (I_K - A^T(AA^T)^{-1}A)w$$

   is a solution to $Ax = y$. The set of solutions $N = \{x \in \mathbb{R}^K : Ax = 0\}$ is a subspace of $\mathbb{R}^K$ because $0 \in N$ and if $x_1, x_2 \in N$ then $\alpha_1 x_1 + \alpha_2 x_2 \in N$. In fact, N is of dimension $K - n$. Now consider the general case with $y \neq 0$.

   Then for any $K \times 1$ vector w the vector

   $$x = (I_K - A^T(AA^T)^{-1}A)w + A^T(AA^T)^{-1}y$$

   is a solution to $Ax = y$

3. Suppose that $n > K$. In this case, the are no solutions to $Ax = y$ except in trivial cases. In that case, the best we can hope for is to find a vector x that minimizes the error, i.e.,

   $$x = \arg\min_{x \in \mathbb{R}^K} ||Ax - y|| = \arg\min_{x \in \mathbb{R}^K} (Ax - y)^T (Ax - y).$$

The **Projection theorem** is a famous result of convex analysis that gives the conditions under which there is a unique solution to the minimization problem and characterizes the solution.

Theorem 16.8.

Let $x \in \mathbb{R}^n$ and let L be a subspace of $\mathbb{R}^n$. Then there exists a unique point $\hat{y} \in L$ for which Euclidean distance

$$||x - y||^2 = \sum_{i=1}^{n} (x_i - y_i)^2$$

is minimized over L. In that case, a necessary and sufficient condition for $\hat{y}$ is that the vector $x - \hat{y}$ be orthogonal to L, meaning for any $y \in L$

$$\langle y, x - \hat{y} \rangle = y^T (x - \hat{y}) = \sum_{i=1}^{n} y_i (x_i - \hat{y}_i) = 0$$

Example 16.14.

Let $n = 3$ and

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then $C(X)$ is the set of all vectors in $\mathbb{R}^3$ with third component zero. What is the closest point in $C(X)$ to y? This is

$$(1, 1, 0)^T = \hat{y}, \quad y - \hat{y} = (0, 0, 1)^T$$

In fact $y - \hat{y}$ is orthogonal to $C(X)$ (i.e., to X and any linear combination thereof), i.e., $y - \hat{y} \in C^{\perp}(X) = \{(0, 0, \alpha)^T, \alpha \in \mathbb{R}\}$.

Exercises

Exercise 1

Solution 1

Given matrices $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix}$, compute $A + B$, AB, BA, A^T, and B^T. Verify that $(AB)^T = B^T A^T$.

Exercise 2

Solution 2

For vectors $x = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ and $y = \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}$, calculate the inner product $x^T y$ and the Euclidean norms $||x||$ and $||y||$. Are x and y orthogonal?

Exercise 3

Solution 3

Find the inverse of the matrix $D = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$. Verify that $DD^{-1} = I$.

Exercise 4

Solution 4

Determine if the following set of vectors is linearly dependent or independent: $x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $x_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $x_3 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.

Exercise 5

Solution 5

Calculate the rank of the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$.

Exercise 6

Solution 6

Compute the trace and determinant of the matrix $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$.

Exercise 7

Solution 7

Show that the set $L = \{x \in \mathbb{R}^2 : x_1 + 2x_2 = 0\}$ is a subspace of $\mathbb{R}^2$.

Exercise 8

Solution 8

Find a basis for the subspace L in Exercise 7. What is the dimension of L?

Exercise 9

Solution 9

Given the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, find its eigenvalues and eigenvectors.

Exercise 10

Solution 10

Verify the Cayley-Hamilton theorem for the matrix A in Exercise 9.

Exercise 11

Solution 11

Show that the eigenvectors corresponding to distinct eigenvalues of the matrix A in Exercise 9 are orthogonal.

Exercise 12

Solution 12

Find the eigendecomposition of the matrix A in Exercise 9. Express A in the form $U\Lambda U^T$, where U is an orthonormal matrix and Λ is a diagonal matrix.

Exercise 13

Solution 13

Determine if the matrix $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ is positive definite, negative definite, or indefinite.

Exercise 14

Solution 14

Given matrices $A = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, determine if $A \geq B$.

Exercise 15

Solution 15

Let X be a random vector with covariance matrix $\Sigma = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}$. Find a matrix A such that $A\Sigma A^T = I$.

Exercise 16

Solution 16 Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find the maximum value of $x^T Ax$ subject to the constraint $x^T x = 1$.

Exercise 17

Solution 17

Find $A^{1/2}$ for the matrix A in Exercise 9.

Exercise 18

Solution 18

Consider the system of equations $Ax = y$ where $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and $y = \begin{pmatrix} 5 \\ 11 \end{pmatrix}$. Find the solution vector x.

Exercise 19

Solution 19 Show that for any matrix A, both $A^T A$ and AA^T are symmetric.

Exercise 20

Solution 20 Find a basis for the null space of matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. What is the dimension of the null space?

Solutions

Solution 1

Exercise 1

Given $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix}$:

- $A + B = \begin{pmatrix} 1+0 & 2+(-1) \\ 3+2 & 4+3 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 5 & 7 \end{pmatrix}$. This follows from the definition of matrix **addition**, described in **Section 16.1**.
- $AB = \begin{pmatrix} 1(0)+2(2) & 1(-1)+2(3) \\ 3(0)+4(2) & 3(-1)+4(3) \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 9 \end{pmatrix}$. This follows from the definition of matrix **multiplication**, described in **Section 16.1**.
- $BA = \begin{pmatrix} 0(1)+(-1)(3) & 0(2)+(-1)(4) \\ 2(1)+3(3) & 2(2)+3(4) \end{pmatrix} = \begin{pmatrix} -3 & -4 \\ 11 & 16 \end{pmatrix}$. This follows from the definition of matrix **multiplication**, described in **Section 16.1**. We observe that $AB \neq BA$. This illustrates that matrix multiplication is not **commutative**.
- $A^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$. This follows from the definition of matrix **transpose**, described in **Section 16.1**.
- $B^T = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix}$. This follows from the definition of matrix **transpose**, described in **Section 16.1**.
- $(AB)^T = \begin{pmatrix} 4 & 8 \\ 5 & 9 \end{pmatrix}$.
- $B^T A^T = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 5 & 9 \end{pmatrix}$. Thus $(AB)^T = B^T A^T$.

Solution 2

Exercise 2

Given $x = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$ and $y = \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix}$:

- $x^T y = (1)(0) + (-2)(4) + (3)(-1) = 0 - 8 - 3 = -11$. This follows from the definition of **inner product**, described in **Section 16.1**.

- $||x|| = \sqrt{(1)^2 + (-2)^2 + (3)^2} = \sqrt{1+4+9} = \sqrt{14}$. This follows from the definition of the **Euclidean norm**, described in **Section 16.1**.

- $||y|| = \sqrt{(0)^2 + (4)^2 + (-1)^2} = \sqrt{0+16+1} = \sqrt{17}$. This follows from the definition of the **Euclidean norm**, described in **Section 16.1**.

- Since $x^T y = -11 \neq 0$, x and y are not orthogonal. This follows from the definition of **orthogonal vectors**, described in **Section 16.1**.

Solution 3

Exercise 3

Given $D = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$, the inverse is $D^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/5 \end{pmatrix}$. This follows from the properties of the **inverse of a diagonal matrix**, described in **Section 16.1**.

Verification: $DD^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$.

Solution 4

Exercise 4

Given $x_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $x_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $x_3 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$. We check for linear dependence by setting

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0.$$

$$\alpha_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This gives us the equations:

$$\begin{cases} \alpha_1 - \alpha_2 = 0 \\ 2\alpha_1 + \alpha_2 + 3\alpha_3 = 0 \end{cases}$$

From the first equation, $\alpha_1 = \alpha_2$. Substituting into the second equation, $2\alpha_1 + \alpha_1 + 3\alpha_3 = 0$, which simplifies to $3\alpha_1 + 3\alpha_3 = 0$. We can choose, for instance, $\alpha_1 = 1$, $\alpha_2 = 1$ and $\alpha_3 = -1$. Since we found a non-trivial solution (not all α_i are zero), the vectors are **linearly dependent**, as defined in **Definition 16.1**.

Solution 5

Exercise 5

Given $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}$. The rank of A is the number of linearly independent rows (or columns). Since the last row is all zeros, and the first two rows are linearly independent, the rank of A is 2. This is consistent with the **Definition 16.3** of **rank**.

Solution 6

Exercise 6

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

- The **trace** of A, $\mathrm{tr}(A) = 1 + 3 = 4$. This follows from the definition of **trace**, provided in **Section 16.1**.
- The **determinant** of A, $\det(A) = (1)(3) - (-1)(2) = 3 + 2 = 5$. This follows from the formula for the **determinant of a** 2×2 **matrix**, provided in **Section 16.1**.

Solution 7

Exercise 7

Let $L = \{x \in \mathbb{R}^2 : x_1 + 2x_2 = 0\}$. To show L is a subspace of $\mathbb{R}^2$, we need to show that for any $x, z \in L$ and any $\alpha, \beta \in \mathbb{R}$, $\alpha x + \beta z \in L$. This follows the definition of a **subspace** given in **Definition 16.4**.

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in L$ and $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in L$. Then $x_1 + 2x_2 = 0$ and $z_1 + 2z_2 = 0$.

Now consider $y = \alpha x + \beta z = \begin{pmatrix} \alpha x_1 + \beta z_1 \\ \alpha x_2 + \beta z_2 \end{pmatrix}$. We have:

$$y_1 + 2y_2 = (\alpha x_1 + \beta z_1) + 2(\alpha x_2 + \beta z_2) = \alpha(x_1 + 2x_2) + \beta(z_1 + 2z_2) = \alpha(0) + \beta(0) = 0.$$

Thus, $y \in L$, and L is a subspace of $\mathbb{R}^2$.

Solution 8

Exercise 8

From Exercise 7, $L = \{x \in \mathbb{R}^2 : x_1 + 2x_2 = 0\}$. We can rewrite the condition as $x_1 = -2x_2$. So, any vector in L can be written as $\begin{pmatrix} -2x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

Thus, a basis for L is the set $\{\begin{pmatrix} -2 \\ 1 \end{pmatrix}\}$. The dimension of L is the number of vectors in the basis, which is 1. This follows the definition of **basis** and **dimension** given in **Definition 16.7**.

Solution 9

Exercise 9

Given $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. To find the eigenvalues, we solve $\det(A - \lambda I) = 0$.

$$\det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0.$$

The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 3$. This is consistent with the method for finding **eigenvalues**, described in **Section 16.1.2**.

For $\lambda_1 = 1$, we solve $(A - I)u = 0$:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow u_1 + u_2 = 0.$$

An eigenvector is $u_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

For $\lambda_2 = 3$, we solve $(A - 3I)u = 0$:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow -u_1 + u_2 = 0.$$

An eigenvector is $u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Solution 10

Exercise 10

The characteristic polynomial of $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ is $p(\lambda) = \lambda^2 - 4\lambda + 3$ (from Exercise 9). The Cayley-Hamilton theorem states that $p(A) = 0$.

$$p(A) = A^2 - 4A + 3I = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - 4 \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} + 3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} - \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix} + \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus, the Cayley-Hamilton theorem holds. The statement of the theorem is given in **Section 16.1.2**.

Solution 11

Exercise 11

From Exercise 9, the eigenvectors are $u_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Their inner product is $u_1^T u_2 = (1)(1) + (-1)(1) = 0$. Thus, the eigenvectors are orthogonal. This is consistent with **Theorem 16.1**, which states that eigenvectors corresponding to distinct eigenvalues are orthogonal.

Solution 12

Exercise 12

From Exercise 9, the eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 3$, and the corresponding eigenvectors are $u_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. We need to normalize the eigenvectors to obtain an orthonormal matrix U.

$$||u_1|| = \sqrt{1^2 + (-1)^2} = \sqrt{2} \Rightarrow \hat{u}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

$$||u_2|| = \sqrt{1^2 + 1^2} = \sqrt{2} \Rightarrow \hat{u}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

Then, $A = U\Lambda U^T$. This decomposition is discussed in **Theorem 16.2**.

Solution 13

Exercise 13

Given $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. To determine the definiteness, we can check the eigenvalues.

$$\det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (3-\lambda)(2-\lambda) - 1 = \lambda^2 - 5\lambda + 5 = 0$$

Using the quadratic formula:

$$\lambda = \frac{-(-5) \pm \sqrt{(-5)^2 - 4(1)(5)}}{2(1)} = \frac{5 \pm \sqrt{5}}{2}$$

Both eigenvalues are positive ($\lambda_1 = \frac{5+\sqrt{5}}{2} > 0$ and $\lambda_2 = \frac{5-\sqrt{5}}{2} > 0$). Thus, the matrix is positive definite. This utilizes the relationship between eigenvalues and definiteness, described in **Section 16.1 after Definition 16.10** stating that a symmetric matrix is positive semi-definite if all eigenvalues are non-negative.

Solution 14

Exercise 14

Given $A = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. We check if $A - B$ is positive semi-definite. $A - B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. We check its eigenvalues. The matrix is psd if and only if all of its eigenvalues are non-negative. The eigenvalues were obtained in Exercise 9 and they are 1 and 3, which are both positive. Thus, $A - B$ is psd, and $A \geq B$. This follows from **Definition 16.11**.

Solution 15

Exercise 15

Given $\Sigma = \begin{pmatrix} 4 & 1 \\ 1 & 9 \end{pmatrix}$. We want to find A such that $A\Sigma A^T = I$. We can take $A = \Sigma^{-1/2}$. We first find the eigendecomposition of $\Sigma = U\Lambda U^T$ where Λ is a diagonal matrix with the eigenvalues and U is an orthogonal matrix ($U^T U = I$) formed by the eigenvectors. A will be then $U\Lambda^{-1/2}U^T$.

The characteristic polynomial is: $(4-\lambda)(9-\lambda) - 1 = \lambda^2 - 13\lambda + 35 = 0$.

The eigenvalues are:

$\lambda = \frac{13 \pm \sqrt{13^2 - 4 \cdot 35}}{2} = \frac{13 \pm \sqrt{29}}{2}$. $\lambda_1 = \frac{13 + \sqrt{29}}{2}$ and $\lambda_2 = \frac{13 - \sqrt{29}}{2}$

$\Lambda = \begin{bmatrix} \frac{13+\sqrt{29}}{2} & 0 \\ 0 & \frac{13-\sqrt{29}}{2} \end{bmatrix}$

The eigenvectors need to be computed and the orthogonal matrix U needs to be formed. The matrix $\Sigma^{-1/2} = A = U\Lambda^{-1/2}U^T$

Solution 16

Exercise 16 The maximum value of $x^T Ax$ subject to $x^T x = 1$ is the largest eigenvalue of A. We can find the eigenvalues of $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ by solving the equation $det(A - \lambda I) = 0$ as follows:

$det(A - \lambda I) = det(\begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}) = (2-\lambda)^2 - 1 = 4 - 4\lambda + \lambda^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3) = 0$
.

Thus, the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 3$. The largest eigenvalue is 3. This method follows directly from **Theorem 16.3**.

Solution 17

Exercise 17 From solution 12:

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

$$\Lambda^{1/2} = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{pmatrix}$$

$$A^{1/2} = U\Lambda^{1/2}U^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \frac{1+\sqrt{3}}{2} & \frac{-1+\sqrt{3}}{2} \\ \frac{-1+\sqrt{3}}{2} & \frac{1+\sqrt{3}}{2} \end{pmatrix}$$

The method for computing $A^{1/2}$ is shown in the paragraph preceding **Theorem 16.4**.

Solution 18

Exercise 18 We have $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, and $y = \begin{pmatrix} 5 \\ 11 \end{pmatrix}$. Since $n = K = 2$ we have the unique solution $x = A^{-1}y$.

We calculate A^{-1}. $det(A) = (1)(4) - (2)(3) = -2$ $A^{-1} = \frac{1}{-2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.

$x = A^{-1}y = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} \begin{pmatrix} 5 \\ 11 \end{pmatrix} = \begin{pmatrix} -10+11 \\ 15/2 - 11/2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Solution 19

[Exercise 19](#) Let A be a matrix. By definition of the transpose operation, $(A^T)^T = A$. Then, $(A^TA)^T = A^T(A^T)^T = A^TA$. So, A^TA is symmetric. Similarly, $(AA^T)^T = (A^T)^TA^T = AA^T$. Therefore, AA^T is also symmetric. This solution uses the definition of the **transpose** operation, described in **Section 16.1**.

Solution 20

[Exercise 20](#)

The null space $N(A)$ of the matrix A is the set of vectors x such that $Ax = 0$. Let $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ Then $Ax = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ The system reduces to one equation $x_1 + 2x_2 = 0$ or $x_1 = -2x_2$. Then, the solution is $x = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Thus a basis for the null space of A is the set $\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\}$ and the dimension of the null space is 1. This follows the definition of **null space** given in **Definition 16.6**, and the definition of **basis** and **dimension** in **Definition 16.7**.

R Script Examples

R Script 1: Matrix Operations and Linear Dependence

```
# Load necessary libraries
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ stringr   1.5.1
✔ ggplot2   3.5.1     ✔ tibble    3.2.1
✔ lubridate 1.9.4     ✔ tidyr     1.3.1
✔ purrr     1.0.2
── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
# Define two matrices
A <- matrix(c(1, 2, 3, 4), nrow = 2, byrow = TRUE)
B <- matrix(c(0, -1, 2, 3), nrow = 2, byrow = TRUE)

# Matrix addition
A_plus_B <- A + B
print("A + B:")
```

```
[1] "A + B:"
```

```
print(A_plus_B)
```

```
     [,1] [,2]
[1,]    1    1
[2,]    5    7
```

```
# Matrix multiplication
AB <- A %*% B
print("A * B:")
```

```
[1] "A * B:"
```

```
print(AB)

     [,1] [,2]
[1,]    4    5
[2,]    8    9

# Matrix transpose
A_transpose <- t(A)
print("A Transpose:")

[1] "A Transpose:"

print(A_transpose)

     [,1] [,2]
[1,]    1    3
[2,]    2    4

# Define vectors
x <- c(1, -2, 3)
y <- c(0, 4, -1)
z <- c(1, 2, 3)

# Inner product
inner_product <- t(x) %*% y
print("Inner Product of x and y:")

[1] "Inner Product of x and y:"

print(inner_product)

     [,1]
[1,]  -11

# Euclidean norm
norm_x <- norm(x, type = "2")  # "2" specifies Euclidean norm
print("Euclidean Norm of x:")

[1] "Euclidean Norm of x:"

print(norm_x)

[1] 3.741657

# Check for linear dependence
# Define a matrix with the vectors as columns
vectors_matrix <- cbind(x, y, z)
#Calculate the determinant
det_vectors_matrix = det(vectors_matrix)
# Vectors are linearly dependent if the determinant of the matrix is 0.
print("Determinant of the matrix formed by the vectors")

[1] "Determinant of the matrix formed by the vectors"

print(det_vectors_matrix)

[1] 4

#Example of linearly dependent vectors
x1 = c(1,1)
x2 = c(-3, -3)
x3 = c(0,0)
```

```
linearly_dependent_matrix = cbind(x1,x2,x3)
print("Determinant of a matrix with linearly dependent columns")
```

```
[1] "Determinant of a matrix with linearly dependent columns"
```

```
print(try(det(linearly_dependent_matrix), TRUE))
```

```
[1] "Error in determinant.matrix(x, logarithm = TRUE, ...) : \n  'x' must be a square matrix\n"
attr(,"class")
[1] "try-error"
attr(,"condition")
<simpleError in determinant.matrix(x, logarithm = TRUE, ...): 'x' must be a square matrix>
```

Explanation:

1. **Load Libraries:** The `tidyverse` library is loaded for general data manipulation and visualization. It includes packages like `dplyr`, `ggplot2`, etc.
2. **Define Matrices:** Two 2x2 matrices, `A` and `B`, are defined using the `matrix()` function. The `nrow` argument specifies the number of rows, and `byrow = TRUE` indicates that the matrix should be filled row-wise. This corresponds to the **definition of a matrix** in Section 16.1.
3. **Matrix Addition:** `A + B` performs element-wise addition of the two matrices. This is consistent with the definition of **matrix addition** in Section 16.1.
4. **Matrix Multiplication:** `A %*% B` performs matrix multiplication. The `%*%` operator is specifically for matrix multiplication in R. This follows the rules of **matrix multiplication** in Section 16.1.
5. **Matrix Transpose:** `t(A)` calculates the transpose of matrix `A`. The rows become columns and vice versa. This corresponds to the **definition of the transpose of a matrix** provided in **Section 16.1**.
6. **Define Vectors:** Two vectors, `x` and `y`, are defined using the `c()` function, which creates vectors. This corresponds to the **definition of vectors** in **Section 16.1**.
7. **Inner Product:** `t(x) %*% y` computes the inner product (dot product) of `x` and `y`. We transpose `x` to make it a row vector, allowing for matrix multiplication with the column vector `y`. This is consistent with the **definition of an inner product** in **Section 16.1**.
8. **Euclidean Norm:** `norm(x, type = "2")` calculates the Euclidean norm (or magnitude) of vector `x`. The `type = "2"` argument specifies the Euclidean norm (L2 norm). This relates to the definition of the **Euclidean norm** in **Section 16.1**.
9. **Check for linear dependence:** We form a matrix whose columns are the vectors we defined. If the determinant of this matrix is zero, this would imply linear dependence. Then we show an example with linearly dependent vectors. This follows from **Definitions 16.1, 16.2 and 16.3**.

R Script 2: Eigenvalues and Eigenvectors

```
# Load necessary libraries
library(tidyverse)

# Define a symmetric matrix
A <- matrix(c(2, 1, 1, 2), nrow = 2, byrow = TRUE)

# Calculate eigenvalues and eigenvectors
eigen_results <- eigen(A)

# Extract eigenvalues
eigenvalues <- eigen_results$values
print("Eigenvalues:")
```

```
[1] "Eigenvalues:"
```

```
print(eigenvalues)
```

```
[1] 3 1
```

```
# Extract eigenvectors
eigenvectors <- eigen_results$vectors
print("Eigenvectors:")
```

```
[1] "Eigenvectors:"
```

```
print(eigenvectors)
```

```
          [,1]       [,2]
[1,] 0.7071068 -0.7071068
[2,] 0.7071068  0.7071068
```

```
# Verify that eigenvectors corresponding to distinct eigenvalues are orthogonal.
#Check that eigenvectors are orthogonal (inner product is approximately zero)
orthogonal_check = t(eigenvectors[,1])%*%eigenvectors[,2]
print("Inner product of eigenvectors")
```

```
[1] "Inner product of eigenvectors"
```

```
print(orthogonal_check)
```

```
             [,1]
[1,] 2.237114e-17
```

```
# Verify the equation A * eigenvector = eigenvalue * eigenvector
# For the first eigenvalue and eigenvector
verification1 <- A %*% eigenvectors[, 1]
print("A * eigenvector1:")
```

```
[1] "A * eigenvector1:"
```

```
print(verification1)
```

```
        [,1]
[1,] 2.12132
[2,] 2.12132
```

```
print("eigenvalue1 * eigenvector1")
```

```
[1] "eigenvalue1 * eigenvector1"
```

```
print(eigenvalues[1] * eigenvectors[, 1])
```

```
[1] 2.12132 2.12132
```

Explanation:

1. **Define a Symmetric Matrix:** A 2x2 symmetric matrix `A` is defined.
2. **Calculate Eigenvalues and Eigenvectors:** The `eigen()` function computes both the eigenvalues and eigenvectors of the matrix `A`. This directly implements the concepts described in **Section 16.1.2**.
3. **Extract Eigenvalues:** `eigen_results$values` extracts the eigenvalues from the results of the `eigen()` function.
4. **Extract Eigenvectors:** `eigen_results$vectors` extracts the eigenvectors. Each column of this matrix represents an eigenvector.
5. **Verify Orthogonality:** The eigenvectors are checked if they are orthogonal, which follows from **Theorem 16.1**.
6. **Verify Eigenvalue-Eigenvector Relationship:** We verify the fundamental relationship $Au = \lambda u$, where u is an eigenvector and λ is the corresponding eigenvalue. We check this relationship by explicitly multiplying the eigenvector by the matrix A and multiplying the eigenvector by the eigenvalue.

R Script 3: Eigendecomposition and Positive Definiteness

```
# Load necessary libraries
library(tidyverse)

# Define a symmetric matrix
A <- matrix(c(3, 1, 1, 2), nrow = 2, byrow = TRUE)

# Calculate eigenvalues and eigenvectors
eigen_results <- eigen(A)
eigenvalues <- eigen_results$values
eigenvectors <- eigen_results$vectors

# Eigendecomposition: A = U * Lambda * U^T
U <- eigenvectors
Lambda <- diag(eigenvalues)  # Create a diagonal matrix from eigenvalues
U_transpose <- t(U)

# Reconstruct A
A_reconstructed <- U %*% Lambda %*% U_transpose
print("Reconstructed A:")
```

```
[1] "Reconstructed A:"
```

```
print(A_reconstructed)
```

```
     [,1] [,2]
[1,]    3    1
[2,]    1    2
```

```
# Check for positive definiteness
# A matrix is positive definite if all its eigenvalues are positive.
is_positive_definite <- all(eigenvalues > 0)
print("Is A positive definite?")
```

```
[1] "Is A positive definite?"
```

```
print(is_positive_definite)
```

```
[1] TRUE
```

```
#Calculate the square root of A
Lambda_sqrt = diag(sqrt(eigenvalues))
A_sqrt = U %*% Lambda_sqrt %*% U_transpose
print("Square root of A")
```

```
[1] "Square root of A"
```

```
print(A_sqrt)
```

```
          [,1]      [,2]
[1,] 1.7013016 0.3249197
[2,] 0.3249197 1.3763819
```

```
#Check that A_sqrt %*% A_sqrt equals approximately A
print("A_sqrt %*% A_sqrt")
```

```
[1] "A_sqrt %*% A_sqrt"
```

```
print(A_sqrt %*% A_sqrt)
```

```
     [,1] [,2]
[1,]    3    1
[2,]    1    2
```

Explanation:

1. **Define a Symmetric Matrix:** A 2x2 symmetric matrix `A` is defined.
2. **Calculate Eigenvalues and Eigenvectors:** The `eigen()` function is used to compute the eigenvalues and eigenvectors.
3. **Eigendecomposition:** The matrix `A` is decomposed into $U\Lambda U^T$, where:
   - `U` is the matrix of eigenvectors.
   - `Lambda` is a diagonal matrix with the eigenvalues on the diagonal (created using `diag(eigenvalues)`).
   - `U_transpose` is the transpose of `U`. This follows **Theorem 16.2**.
4. **Reconstruct A:** The matrix `A` is reconstructed from its eigendecomposition to verify the decomposition is correct.
5. **Check for Positive Definiteness:** The `all(eigenvalues > 0)` checks if all eigenvalues are strictly positive. If true, the matrix is positive definite, as discussed after **Definition 16.10**.
6. **Calculate the square root of A:** We take the square root of the eigenvalues and then reconstruct the matrix $A^{1/2}$. This implements the method described in the text before **Theorem 16.4**.

R Script 4: Covariance Matrix and Standardization

```
# Load necessary libraries
library(tidyverse)

# Simulate data
set.seed(123) # for reproducibility
n <- 100 # number of observations
x1 <- rnorm(n, mean = 2, sd = 1)
x2 <- rnorm(n, mean = -1, sd = 2)
X <- cbind(x1, x2)

# Calculate the sample covariance matrix
sample_covariance <- cov(X)
print("Sample Covariance Matrix:")
```

```
[1] "Sample Covariance Matrix:"
```

```
print(sample_covariance)
```

```
            x1          x2
x1  0.83323283 -0.08744214
x2 -0.08744214  3.74025239
```

```
# Calculate eigenvalues and eigenvectors
eigen_results <- eigen(sample_covariance)
eigenvalues <- eigen_results$values
eigenvectors <- eigen_results$vectors

# Calculate Sigma^(-1/2)
Sigma_inv_sqrt <- eigenvectors %*% diag(1/sqrt(eigenvalues)) %*% t(eigenvectors)
print("Sigma^(-1/2):")
```

```
[1] "Sigma^(-1/2):"
```

```
print(Sigma_inv_sqrt)
```

```
           [,1]       [,2]
[1,] 1.09671911 0.01742535
```

```
[2,] 0.01742535 0.51741237
```

```
# Standardize the data
X_centered <- scale(X, center = TRUE, scale = FALSE)  # Center the data
Z <- X_centered %*% Sigma_inv_sqrt
print("First few rows of the standardized data")
```

```
[1] "First few rows of the standardized data"
```

```
print(head(Z))
```

```
            [,1]        [,2]
[1,] -0.73484432 -0.63519604
[2,] -0.33888928  0.37153543
[3,]  1.60546601 -0.11840509
[4,] -0.03018601 -0.24870031
[5,]  0.01322596 -0.87278882
[6,]  1.78397350  0.09300653
```

```
# Verify that the covariance matrix of Z is approximately the identity matrix
cov_Z <- cov(Z)
print("Covariance Matrix of Z:")
```

```
[1] "Covariance Matrix of Z:"
```

```
print(cov_Z)
```

```
              [,1]          [,2]
[1,]  1.000000e+00 -1.121437e-17
[2,] -1.121437e-17  1.000000e+00
```

Explanation:

1. **Simulate Data:** Two random variables, `x1` and `x2`, are simulated using `rnorm()` (normally distributed). These are combined into a matrix `X`.
2. **Calculate Sample Covariance Matrix:** The `cov(X)` function computes the sample covariance matrix of `X`. This relates to **Definition 16.12**.
3. **Calculate Eigenvalues and Eigenvectors:** The `eigen()` function is used to find the eigenvalues and eigenvectors of the covariance matrix.
4. **Calculate $\Sigma^{-1/2}$:** The matrix $\Sigma^{-1/2}$ is calculated using the eigendecomposition. The square root of the inverse of each eigenvalue is taken (since we want $\Sigma^{-1/2}$), and these are placed on the diagonal of a matrix. This follows the description of $\Sigma^{-1/2}$ and **Theorem 16.5**.
5. **Standardize the Data:**
   - `X_centered <- scale(X, center = TRUE, scale = FALSE)` centers the data by subtracting the mean of each column.
   - `Z <- X_centered %*% Sigma_inv_sqrt` multiplies the centered data by $\Sigma^{-1/2}$ to obtain the standardized data `Z`. This implements the standardization process from **Theorem 16.5**.
6. **Verify Covariance of Z:** `cov(Z)` calculates the covariance matrix of the standardized data `Z`. This should be approximately equal to the identity matrix, as stated in **Theorem 16.5**.

R Script 5: System of Linear Equations and Projection

```
# Load necessary libraries
library(tidyverse)

# Define matrix A and vector y
A <- matrix(c(1, 2, 3, 4), nrow = 2, byrow = TRUE)
y <- c(5, 11)
```

```
# Solve the system Ax = y
x <- solve(A, y)
print("Solution x:")
```

```
[1] "Solution x:"
```

```
print(x)
```

```
[1] 1 2
```

```
# Projection Example (n > K)
# Define matrix X and vector y (n=3, K=2)
X <- matrix(c(1, 1, 0, 1, 0, 0), nrow = 3, byrow=FALSE)
y_proj <- c(1, 1, 1)

# Find projection using the formula:  x_hat = (X^T X)^(-1) X^T y
X_transpose <- t(X)
x_hat <- solve(t(X) %*% X) %*% t(X) %*% y_proj
print("x_hat")
```

```
[1] "x_hat"
```

```
print(x_hat)
```

```
     [,1]
[1,]    1
[2,]    0
```

```
#Calculate the projection y_hat = X %*% x_hat
y_hat = X %*% x_hat
print("Projection of y onto C(X):")
```

```
[1] "Projection of y onto C(X):"
```

```
print(y_hat)
```

```
     [,1]
[1,]    1
[2,]    1
[3,]    0
```

```
# Check orthogonality: (y - y_hat) should be orthogonal to the columns of X
orthogonality_check <- t(X) %*% (y_proj - y_hat)
print("Orthogonality Check:")
```

```
[1] "Orthogonality Check:"
```

```
print(orthogonality_check)
```

```
     [,1]
[1,]    0
[2,]    0
```

Explanation:

1. **Define Matrix and Vector:** A 2x2 matrix `A` and a 2x1 vector `y` are defined. This sets up a system of linear equations $Ax = y$, as described in **Section 16.2**.
2. **Solve the System:** The `solve(A, y)` function solves the system of equations. This corresponds to the case $n = K$ in **Section 16.2**, where a unique solution exists if `A` is invertible.
3. **Projection Example** Another matrix X and vector y are defined, with the number of rows of matrix X being greater than the number of its columns.

4. **Find projection**: We implement the formula $x = (X^T X)^{-1} X^T y$ to find the projection of vector y onto the column space of matrix X.
5. **Calculate y_hat:** `y_hat` is the projection of `y` onto the column space of `X`.
6. **Check Orthogonality:** We check that the difference between `y` and its projection `y_hat` $(y - \hat{y})$ is orthogonal to the columns of X. This verification confirms the properties of the projection as stated in **Theorem 16.8**. The inner product of $(y - \hat{y})$ and each column of X should be approximately zero.

YouTube Videos for Linear Algebra Concepts

Here are some YouTube videos that explain the concepts mentioned in the attached text. I have verified that all videos are currently available (as of October 26, 2023).

1. Matrices, Vectors, Transpose, Matrix Addition, and Multiplication

- **Title:** Essence of linear algebra, Chapter 1
- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=fNk_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab
- **Relation to Text:** This video provides a strong visual and intuitive introduction to vectors and matrices, covering the basics of what they represent geometrically. It gives context to operations like matrix addition, scalar multiplication which are defined in **Section 16.1**. It complements the text's algebraic definitions with geometric interpretations.
- **Title:** Matrices - Intro (Part 1) | Don't Memorise
- **Channel:** Infinity Learn NEET
- **URL**: https://www.youtube.com/watch?v=MxJz3AHjK8k
- **Relation to Text**: This video presents an introduction to the very basics of matrix algebra. It covers the representation of a matrix, and the definition of basic operations such as addition, substraction, multiplication. It is a great complement of the formal definition presented in **Section 16.1**.
- **Title:** Matrix multiplication as composition | Essence of linear algebra, Chapter 4
- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=XkY2DOUCWMU
- **Relation to Text:** This video focuses specifically on matrix multiplication, explaining it as a composition of linear transformations. This provides a deeper understanding of why matrix multiplication is defined the way it is (and why it's not commutative), complementing the algebraic definition in **Section 16.1**.

2. Inner Product, Norm, Orthogonality

- **Title:** Dot products and duality | Essence of linear algebra, Chapter 9
- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=LyGKycYT2v0

- **Relation to Text:** This video explains the dot product (inner product) geometrically and connects it to the concept of duality. This complements the definition of the **inner product** in **Section 16.1** and the definition of **orthogonal vectors**. It also helps visualize the **Euclidean norm**, which is defined using the inner product.

- **Title:** Vector Dot Product and Vector Length

- **Channel:** Khan Academy

- **URL:** https://www.youtube.com/watch?v=QykUMjWAWRw

- **Relation to text:** This video covers the concept of **norm** and **inner product**, which are discussed in **Section 16.1**

3. Linear Dependence, Independence, Rank, Span, Basis

- **Title:** Linear combinations, span, and basis vectors | Essence of linear algebra, Chapter 2

- **Channel:** 3Blue1Brown

- **URL:** https://www.youtube.com/watch?v=k7RM-ot2NWY

- **Relation to Text:** This video introduces the concepts of **linear combinations**, **span**, and **basis vectors**, providing a visual and intuitive understanding. This directly relates to **Definitions 16.1, 16.2, 16.5, and 16.7**.

- **Title:** Linear independence and linear dependence, examples

- **Channel:** blackpenredpen

- **URL:** https://www.youtube.com/watch?v=c_g4Z-M4-9U

- **Relation to Text:** This video uses concrete examples to describe **linear dependence** and **independence**. It complements **Definition 16.1** and **Definition 16.2**.

- **Title:** Introduction to the Rank of a Matrix

- **Channel:** Mathispower4u

- **URL:** https://www.youtube.com/watch?v=85kLqWIS_dI

- **Relation to text:** The video provides a short and useful definition of matrix **rank** and its properties. It complements **Definition 16.3**.

4. Trace and Determinant

- **Title:** The Trace of a Matrix

- **Channel:** TheTrevTutor

- **URL:** https://www.youtube.com/watch?v=-iz8kl_nagc

- **Relation to Text:** This video provides a good introduction to the concept of **trace of a matrix**, and complements the definition in **Section 16.1**.

- **Title:** Essence of linear algebra, Chapter 6

- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=Ip3X9LOh2dk
- **Relation to Text:** This video explains the **determinant** geometrically, as the factor by which a linear transformation changes areas or volumes. This provides intuition behind the determinant formula provided in **Section 16.1**.

5. Inverse, Linear Spaces, Null Space

- **Title:** Inverse matrices, column space and null space | Essence of linear algebra, Chapter 7
- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=uQhTuRlWMxw
- **Relation to Text:** This video covers **inverse matrices**, **column space**, and **null space**, connecting them to solving systems of linear equations. This relates directly to the discussion of inverses in **Section 16.1**, **Definition 16.6** (null space), and the introduction of **linear spaces** in **Section 16.1.1**.

6. Eigenvectors and Eigenvalues

- **Title:** Essence of linear algebra, Chapter 14
- **Channel:** 3Blue1Brown
- **URL:** https://www.youtube.com/watch?v=PFDu9oVAE-g
- **Relation to Text:** This video provides a visual and intuitive explanation of **eigenvectors** and **eigenvalues**, explaining them as vectors that are only scaled (not rotated) by a linear transformation. This directly corresponds to **Definition 16.9** and the surrounding discussion in **Section 16.1.2**.
- **Title:** Eigenvectors and eigenvalues | MIT 18.06SC Linear Algebra, Fall 2011
- **Channel:** MIT OpenCourseWare
- **URL:** https://www.youtube.com/watch?v=ue-1gfI6m4Y
- **Relation to text:** This video complements the information about **eigenvalues** and **eigenvectors** provided in **Section 16.1.2**.
- **Title:** The Applications of Eigenvectors & Eigenvalues | How to find and apply them
- **Channel:** Zach Star
- **URL:** https://www.youtube.com/watch?v=-Z2xO-a7zIQ
- **Relation to text:** This video gives another perspective on the topic of **eigenvalues** and **eigenvectors**, and also discusses applications, which is aligned with **Section 16.1.3** in the text.

7. Eigendecomposition, Positive Definite Matrices

- **Title:** [Linear Algebra] Eigendecomposition
- **Channel:** Jacobo Sarin
- **URL:** https://www.youtube.com/watch?v=EOtB7D681U8
- **Relation to Text**: This video describes **eigendecomposition** and its properties, complementing **Theorem 16.2**.
- **Title:** Definite Matrices

- **Channel** : Dr. Trefor Bazett
- **URL**: https://www.youtube.com/watch?v=ZQtfmJHGQkI
- **Relation to Text**: This video presents a clear definition of **positive definite matrices** and some of their properties. This relates to **Definition 16.10**.

8. Systems of Linear Equations and Projections

- **Title:** Solving Ax = b | MIT 18.06SC Linear Algebra, Fall 2011
- **Channel:** MIT OpenCourseWare
- **URL:** https://www.youtube.com/watch?v=9s_0rAiwg78
- **Relation to Text:** This lecture discusses solving systems of linear equations, which is the primary topic of **Section 16.2**.
- **Title:** Visualizing Projections | MIT 18.06SC Linear Algebra, Fall 2011
- **Channel**: MIT OpenCourseWare
- **URL:**https://www.youtube.com/watch?v=qczPZp7Z-b4
- **Relation to Text:** This lecture explains projections visually, which is directly related to **Theorem 16.8** and the example following it.

Multiple Choice Exercises

MC Exercise 1

MC Solution 1

What is the transpose of the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$?

a. $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$
b. $\begin{pmatrix} 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$
c. $\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$
d. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

MC Exercise 2

MC Solution 2

If A is an $n \times m$ matrix and B is a $p \times q$ matrix, under what condition can the matrix product AB be defined?

a. $n = p$
b. $m = p$

c. $n = q$
d. $m = q$

MC Exercise 3

MC Solution 3

Which of the following properties holds for the trace of matrices (assuming the matrices are conformable)?

a. $\mathrm{tr}(AB) = \mathrm{tr}(A)\mathrm{tr}(B)$
b. $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$
c. $\mathrm{tr}(A^T) = 1/\mathrm{tr}(A)$
d. $\mathrm{tr}(AB) = \mathrm{tr}(A)\ /\ \mathrm{tr}(B)$

MC Exercise 4

MC Solution 4

What is the inner product of the vectors $x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $y = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$?

a. 4
b. 5
c. 6
d. 7

MC Exercise 5

MC Solution 5

What is the Euclidean norm of the vector $x = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$?

a. 1
b. 5
c. 7
d. 25

MC Exercise 6

MC Solution 6

Two vectors x and y are orthogonal if:

a. $x^T y = 1$
b. $x^T y = 0$
c. $||x|| = ||y||$
d. $x = y$

MC Exercise 7

MC Solution 7

A set of vectors is linearly dependent if:

a. All the vectors are identical.
b. The only linear combination that equals the zero vector is the trivial one (all coefficients are zero).
c. There exists a non-trivial linear combination of the vectors that equals the zero vector.
d. The vectors are all orthogonal to each other.

MC Exercise 8

MC Solution 8

What is the determinant of the matrix $\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}$?

a. 2
b. 5
c. 10
d. 14

MC Exercise 9

MC Solution 9

If A is an $n \times n$ matrix, which of the following statements is true about the determinant of A?

a. $\det(A^T) = 1/\det(A)$
b. $\det(A^T) = -\det(A)$
c. $\det(A^T) = \det(A)$
d. $\det(A^T) = n \cdot \det(A)$

MC Exercise 10

MC Solution 10

A square matrix A is invertible (nonsingular) if:

a. Its determinant is zero.
b. Its rank is less than its dimension.
c. Its determinant is non-zero.
d. It is a diagonal matrix.

MC Exercise 11

MC Solution 11

What is the inverse of the matrix $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, where a and b are non-zero?

a. $\begin{pmatrix} b & 0 \\ 0 & a \end{pmatrix}$
b. $\begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix}$

c. $\begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix}$
d. $\begin{pmatrix} 0 & 1/b \\ 1/a & 0 \end{pmatrix}$

MC Exercise 12

MC Solution 12

An eigenvector u of a matrix A satisfies which of the following equations?

a. $Au = 0$
b. $Au = u$
c. $Au = \lambda u$ for some scalar λ
d. $A^T u = Au$

MC Exercise 13

MC Solution 13

If a matrix A has an eigenvalue λ, what is the corresponding eigenvalue of $A + 2I$, where I is the identity matrix?

a. λ
b. $\lambda + 2$
c. 2λ
d. λ^2

MC Exercise 14

MC Solution 14

A real symmetric matrix is positive definite if:

a. All its entries are positive.
b. Its determinant is positive.
c. All its eigenvalues are positive.
d. All its eigenvalues are non-negative.

MC Exercise 15

MC Solution 15

The rank of a matrix is:

a. The number of its rows.
b. The number of its columns.
c. The number of its linearly independent rows (or columns).
d. The sum of its diagonal elements.

MC Exercise 16

MC Solution 16

Which of the following is a property of the eigenvalues of a real symmetric matrix?

a. They can be complex.
b. They are always real.
c. They are always positive.
d. They are always zero.

MC Exercise 17

MC Solution 17

A set of vectors is called a basis for a vector space if: (a) The vectors are linearly dependent. (b) The vectors span the vector space and are linearly independent. (c) The vectors are orthogonal. (d) The vectors have unit length.

MC Exercise 18

MC Solution 18 If X is a random variable with $E(X) = \mu$, what does $Cov(X)$ represent? (a) $E(X^2) - \mu$ (b) $E(XX^T) - \mu\mu^T$ (c) $E(X) - \mu^2$ (d) $E((X - \mu)^2)$

MC Exercise 19

MC Solution 19 If a matrix A is singular, it means that: (a) A is not a square matrix. (b) A^{-1} exists. (c) det(A) = 0. (d) All elements of A are zero.

MC Exercise 20

MC Solution 20

The dimension of the null space of a matrix A is also known as the:

a. Rank of A.
b. Nullity of A.
c. Trace of A.
d. Determinant of A. ## Multiple Choice Solutions

MC Solution 1

MC Exercise 1

The correct answer is (a). The **transpose** of a matrix is obtained by interchanging its rows and columns. This is described in **Section 16.1**.

MC Solution 2

MC Exercise 2

The correct answer is (b). For the matrix product AB to be defined, the number of columns of A (which is m) must be equal to the number of rows of B (which is p). This condition is stated in the definition of **matrix multiplication** in **Section 16.1**.

MC Solution 3

MC Exercise 3

The correct answer is (b). The trace of the sum of two matrices is the sum of their traces. This is one of the properties of the **trace** operation stated in **Section 16.1**.

MC Solution 4

MC Exercise 4

The correct answer is (b). The **inner product** is calculated as: $x^T y = (1)(-1) + (2)(3) = -1 + 6 = 5$. This calculation is shown in **Section 16.1**.

MC Solution 5

MC Exercise 5

The correct answer is (b). The **Euclidean norm** is calculated as: $||x|| = \sqrt{(3)^2 + (-4)^2} = \sqrt{9 + 16} = \sqrt{25} = 5$. The formula for the **norm** is defined in **Section 16.1**.

MC Solution 6

MC Exercise 6

The correct answer is (b). Two vectors are **orthogonal** if their inner product is zero. This definition is provided in **Section 16.1**.

MC Solution 7

MC Exercise 7

The correct answer is (c). This is the definition of **linear dependence** given in **Definition 16.1**.

MC Solution 8

MC Exercise 8

The correct answer is (b). The **determinant** is calculated as: $(2)(4) - (1)(3) = 8 - 3 = 5$. This follows the formula given in **Section 16.1**.

MC Solution 9

MC Exercise 9

The correct answer is (c). The determinant of the transpose of a matrix is equal to the determinant of the original matrix.

MC Solution 10

MC Exercise 10

The correct answer is (c). A square matrix is **invertible** (nonsingular) if and only if its determinant is non-zero. This fact is mentioned in **Section 16.1** when discussing the inverse of a matrix.

MC Solution 11

[MC Exercise 11](#)

The correct answer is (b). This is a special case of the **inverse of a diagonal matrix**, discussed in **Section 16.1**.

MC Solution 12

[MC Exercise 12](#)

The correct answer is (c). This is the definition of an **eigenvector** and its corresponding **eigenvalue**, as given in **Definition 16.9**.

MC Solution 13

[MC Exercise 13](#)

The correct answer is (b). If $Au = \lambda u$, then $(A + 2I)u = Au + 2Iu = \lambda u + 2u = (\lambda + 2)u$. This property is discussed in the text after **Theorem 16.2**.

MC Solution 14

[MC Exercise 14](#)

The correct answer is (c). A real symmetric matrix is **positive definite** if all its eigenvalues are positive. This condition is explained in **Section 16.1** after **Definition 16.10**. Option (d) describes a **positive semi-definite** matrix.

MC Solution 15

[MC Exercise 15](#)

The correct answer is (c). This is the definition of the **rank** of a matrix, given in **Definition 16.3**.

MC Solution 16

[MC Exercise 16](#)

The correct answer is (b). The **eigenvalues** of a real symmetric matrix are always real. This is mentioned in **Section 16.1.2**.

MC Solution 17

[MC Exercise 17](#)

The correct answer is (b). This is the definition of a **basis** for a vector space given in **Definition 16.7**.

MC Solution 18

[MC Exercise 18](#)

The correct answer is (b). $Cov(X)$ is defined in **Definition 16.12** as $E[(X - \mu)(X - \mu)^T] = E[XX^T] - \mu\mu^T$.

MC Solution 19

MC Exercise 19

The correct answer is (c). A matrix A is **singular** if its determinant is 0. This fact is related to the discussion of **inverse** matrices in **Section 16.1**.

MC Solution 20

MC Exercise 20 The correct answer is (b). The dimension of the null space of A is called **nullity** of A. The definition of the **null space** of a matrix appears in **Definition 16.6**.

Author: Peter Fuleky