Tutorial 2: Review of Linear Algebra

Mohammadreza Safavi
mohammadreza.safavi@mail.utoronto.ca

University of Toronto
ECE1513: Introduction to Machine Learning
Instructor: Dr. Ali Bereyhi

Jan 20, 2025

Linear algebra

We will review some basic concepts in linear algebra.

- Vectors
- Matrices
- Dot product
- Cross product
- Matrix-vector multiplication
- Basis vectors
- Change of basis
- Eigenvalues and eigenvectors
- Singular Value Decomposition

Vectors

A vector x is a one-dimensional array of numbers. By convention, we represent vectors as column vectors. e.g.:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

Matrices

A matrix A is a two-dimensional array of numbers. By convention, we represent matrices as:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{bmatrix}$$

Inner product

The inner product of two n-dimensional vectors x and y, is defined as:

$$x \cdot y = x^T y = y^T x = \sum_{i=1}^{N} x_i y_i = [x_1, \ldots, u_N] \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

Cross product

The outer product of two n-dimensional vectors x and y, is defined as:

$$x \times y = xy^T = (yx^T)^T = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \begin{bmatrix} y_1, \ldots, y_N \end{bmatrix}$$

$$= \begin{bmatrix} x_1 y_1 & \ldots & x_1 y_N \\ \vdots & \ddots & \vdots \\ x_N y_1 & \ldots & x_N y_N \end{bmatrix}$$

Matrix-vector multiplication

Let $A \in R^{N \times M}$ and $x \in R^{M}$, then Ax is defined as:

$$Ax = \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \dots & a_{NM} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} = [a_1, \dots, a_M] \begin{bmatrix} x_1 \\ \vdots \\ x_M \end{bmatrix} = \sum_{i=1}^{M} x_i a_i \in R^{N}$$

Where a_i is the i^{th} column of A.

Matrix-vector multiplication

Alternatively, we can write Ax as:

$$Ax = \begin{bmatrix} a_1^T \\ \vdots \\ a_N^T \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ \vdots \\ a_N^T x \end{bmatrix}$$

Where a_i^T is the i^{th} row of A.

Matrix-matrix multiplication

Let $A \in R^{N \times M}$ and $B \in R^{M \times P}$, then AB is defined as:

$$AB = \begin{bmatrix} a_1, \ldots, a_M \end{bmatrix} \begin{bmatrix} b_1^T \\ \vdots \\ b_P^T \end{bmatrix} = \sum_{i=1}^{M} a_i b_i^T \in R^{N \times P}$$

Where a_i is the i^{th} column of A and b_i^T is the i^{th} row of B.

Matrix-matrix multiplication

Alternatively, we can write AB as:

$$AB = \begin{bmatrix} a_1^T \\ \vdots \\ a_N^T \end{bmatrix} B = \begin{bmatrix} a_1^T \\ \vdots \\ a_N^T \end{bmatrix} \begin{bmatrix} b_1 & \dots & b_P \end{bmatrix} = \begin{bmatrix} a_1^T b_1 & \dots & a_1^T b_P \\ \vdots & \ddots & \vdots \\ a_N^T b_1 & \dots & a_N^T b_P \end{bmatrix} = \begin{bmatrix} a_1^T B \\ \vdots \\ a_N^T B \end{bmatrix}$$

Where a_i^T is the i^{th} row of A and b_j is the j^{th} column of B.

Matrix-matrix multiplication

In another form, we can write AB as:

$$AB = A \begin{bmatrix} b_1 & \dots & b_P \end{bmatrix} = \begin{bmatrix} Ab_1 & \dots & Ab_P \end{bmatrix}$$

Matrix-matrix multiplication

Lastly, we can compute each element of AB using:

$$(AB)_{ij} = \sum_{k=1}^{M} a_{ik} b_{kj}$$

Matrix inversion

The inverse of a square matrix A, denoted as A^{-1}, is defined as:

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

Where I is the identity matrix.

Matrix inversion

For a 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse is:

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

Basis vectors

The basis vectors are the vectors that span the space. The standard basis vectors are:

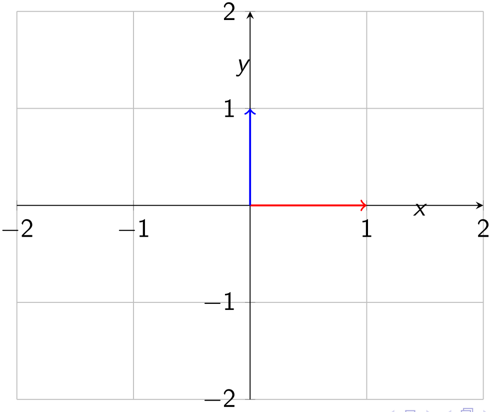
$$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Where the 1 is at the i^{th} position.

Basis vectors example

The standard basis vectors in 2D are:

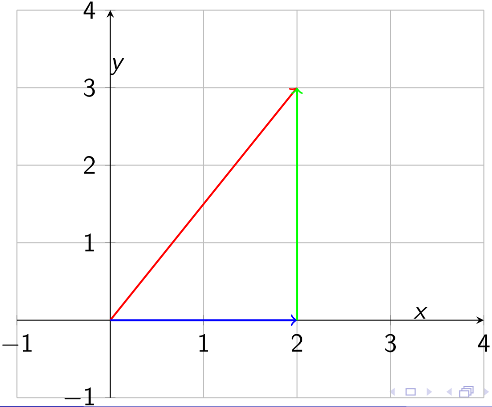
$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



Basis vectors example

As an example, the vector $x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ can be represented as:

$$x = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



Change of basis

But we can also represent x in another basis B. Suppose that we have:

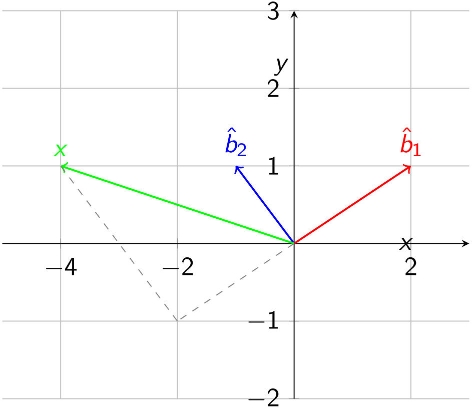
$$\hat{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \hat{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Then if we take a linear combination of $\hat{b}_1$ and $\hat{b}_2$, we can represent a point as:

$$x = -1\hat{b}_1 + 2\hat{b}_2 = -1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

Change of basis

$$x = -1\begin{bmatrix}2\\1\end{bmatrix} + 2\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}-4\\1\end{bmatrix}$$



Change of basis

In order to change the basis of a vector, we can use a matrix A that transforms the vector from basis B to basis C.

Say we put the basis vectors in a matrix B:

$$B = \begin{bmatrix} \hat{b}_1 & \hat{b}_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

Suppose that a vector x is represented in B as:

$$x_B = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

meaning that x is -1 times the first basis vector and 2 times the second basis vector.

Then the vector x can be transformed back to the original basis as:

$$x = Bx_B$$

Here, $x_B = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ so $x = Bx_B = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$.

Change of basis

You can view this as a transformation of the vector $[-1, 2]^T$ into $[-4, 1]^T$ using a matrix B. I.e. $x = Bx_B$ is a linear transformation.

Linear transformation

A linear transformation T is a function that maps a vector x to another vector y such that:

$$T(x) = Ax$$

Where A is a matrix.
Suppose that $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$, then $T(x) = Ax$ is a linear transformation. Say $x = \hat{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $T(\hat{i}) = A\hat{i} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$.
Interestingly, $T(\hat{i})$ is in the same direction as $\hat{i}$, but scaled by a factor of 3!

Linear transformation

Now say we have $x = \hat{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $T(\hat{j}) = A\hat{j} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

This time $T(\hat{j})$ is not in the same direction as $\hat{j}$, i.e., it has been rotated and scaled.

Linear transformation

Now say we have $x = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, then $T(x) = Ax = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
Again, $T(x)$ is in the same direction as x, but scaled by a factor of 2. This means that all the vectors in the direction of x will only be scaled after applying the transformation and they won't be rotated. These vectors are called eigenvectors for the matrix A.

Eigenvalues and eigenvectors

The eigenvectors of a matrix A are the vectors that are only scaled by a factor after applying the transformation $T(x) = Ax$. The scaling factor is called the eigenvalue. The eigenvectors and eigenvalues of a matrix A are the solutions to the equation:

$$Av = \lambda v$$

Where v is the eigenvector and λ is the eigenvalue.

Eigenvalues and eigenvectors

Rewriting the equation $Av = \lambda v$ gives:

$$(A - \lambda I)v = 0$$

Where I is the identity matrix. This equation has a non-trivial solution if the determinant of the matrix $A - \lambda I$ is zero. This gives the characteristic equation:

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

Eigenvalues and eigenvectors

Suppose that we have a matrix $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$. The characteristic equation is:

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) = (3-\lambda)(2-\lambda) = 0$$

Solving this equation gives $\lambda = 3, 2$.

Eigenvalues and eigenvectors

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) = (3-\lambda)(2-\lambda) = 0$$

Solving this equation gives $\lambda = 3, 2$. For $\lambda = 3$, the eigenvector is:

$$Av = 3v \rightarrow \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 3 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

This gives $3v_1 + v_2 = 3v_1$ and $2v_2 = 3v_2$. This means that $v_2 = 0$ and $v_1 = 1$. So the eigenvector for $\lambda = 3$ is $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Eigenvalues and eigenvectors

For $\lambda = 2$, the eigenvector is:

$$Av = 2v \rightarrow \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 2 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

This gives $3v_1 + v_2 = 2v_1$ and $2v_2 = 2v_2$. This means that v_2 could be any value, say $v_2 = 1$ so $3v_1 + 1 = 2v_1$. This gives $v_1 = -1$ and $v_2 = 1$. So the eigenvector for $\lambda = 2$ is $v = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. In practice, we normalize the eigenvectors to have a unit length. So the eigenvector is $v = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

Eigenvalues and eigenvectors

The eigenvectors and eigenvalues of the matrix $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ are:

$$\lambda = 3, v = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \lambda = 2, v = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Eigenvalue decomposition

We can put all the eigenvectors in a matrix V and all the eigenvalues in a diagonal matrix Λ. Using the definition of eigenvectors and eigenvalues, we can write:

$$AV = V\Lambda$$

Now (under some conditions), if we multiply both sides by V^{-1}, we get:

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

Eigenvalue decomposition

The eigenvalue decomposition of a matrix A is:

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

Where V is a matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues. E.g., for a 2x2 matrix A:

$$A = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{-1}$$

Eigenvalue decomposition

For a symmetric matrix A, the eigenvalues are real and the eigenvectors are orthogonal. This means that the matrix A can be decomposed as:

$$A = V \Lambda V^T$$

Where V is a matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues.

Singular Value Decomposition

In certain applications, we need to represent a matrix using lower rank matrices. Say we have a matrix $A \in R^{N \times M}$, then we want to represent it as:

$$A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

This means that we want to represent A as a sum of rank-1 matrices (each $u_i v_i^T$ is a rank-1 matrix as it is the outer product of two vectors).

Singular Value Decomposition

We can rewrite this summation as a matrix multiplication:

$$A = U\Sigma V^T$$

Where U and V are orthogonal matrices and Σ is a diagonal matrix of singular values.

Singular Value Decomposition

The Singular Value Decomposition (SVD) of a matrix A is:

$$A = U\Sigma V^T$$

Where U and V are orthogonal matrices and Σ is a diagonal matrix of singular values.

Singular Value Decomposition

In order to compute the SVD of a matrix A, we can use the eigenvectors and eigenvalues of $A^T A$ and AA^T.
We can form $A^T A$ as:

$$A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T \Sigma V^T$$

Keep in mind $U^T U = I$, as U is an orthogonal matrix.

Singular Value Decomposition

Now if we define $D = \Sigma^T \Sigma$, then $A^T A = VDV^T$, which reminds us of the eigenvalue decomposition!
This means that the eigenvectors of $A^T A$ are the columns of V and the eigenvalues are the diagonal elements of D.

Singular Value Decomposition

Similarly, we can form AA^T as:

$$AA^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma\Sigma^T U^T$$

If we define $D = \Sigma\Sigma^T$, then $AA^T = UDU^T$. This means that the eigenvectors of AA^T are the columns of U and the eigenvalues are the diagonal elements of D.

Singular Value Decomposition

Example: Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. We can compute the SVD of A as:

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

Singular Value Decomposition

To get the eigenvalues of $A^T A$, we can solve the characteristic equation:

$$\det(A^T A - \lambda I) = 0$$

$$\det\left(\begin{bmatrix} 1-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix}\right) = 0$$

This gives $\lambda = 2, 1, 0$.

Singular Value Decomposition

Following the equation $A^T A v = \lambda v$, we can find the eigenvectors of $A^T A$. The eigenvectors of $A^T A$ are:

$$V = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}$$

The singular values are the square roots of the eigenvalues of $A^T A$, i.e., $\sigma = \sqrt{2}, 1, 0$.

Singular Value Decomposition

The eigenvectors of AA^T are:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Singular Value Decomposition

Putting all together, the SVD of A is:

$$A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

Which if you multiply the matrices, you will get back the original matrix A!