Eigenvectors and Eigenvalues

Jaskirat Kaur

February 16, 2024

Understanding Eigenvectors and Eigenvalues

- Eigenvectors are like the "compass directions" for a matrix transformation. They point in directions that remain consistent even after the transformation is applied.
- Eigenvalues tell us how much an eigenvector is stretched or squished during the transformation. A larger absolute value means more stretching or compressing, while a value of 1 means no change.

Role in Machine Learning

- **Dimensionality Reduction:** In algorithms like Principal Component Analysis (PCA), eigenvectors are used to find the directions (principal components) that capture the most variance in the data. Eigenvalues give the amount of variance captured.
- **Data Compression:** By keeping the eigenvectors associated with the largest eigenvalues (thus the most significant features of the data), we can compress high-dimensional data into a lower-dimensional space with minimal loss of information.
- **Google's PageRank Algorithm:** The web pages' rankings are determined by the eigenvector of the largest eigenvalue in the PageRank matrix
- **Noise Reduction:** In signal processing and feature extraction, removing eigenvectors with small eigenvalues can reduce noise and improve model performance.

Eigenvectors Manipulation

Eigenvalue manipulation

Let A be a matrix with eigenvalue λ. Prove each of the following.

1. λ^2 is an eigenvalue of A^2.
2. λ^{-1} is an eigenvalue of A^{-1}, if A is invertible.
3. $\lambda + 1$ is an eigenvalue of $A + I$.

Approach

- An eigenvalue λ of a matrix A represents a scale factor that stretches a vector v, known as an eigenvector.
- The equation $Av = \lambda v$ shows this relationship where the vector v does not change direction when transformed by A.
- Manipulating A (e.g., squaring or adding identity) will similarly manipulate the eigenvalues, affecting the scale of the eigenvectors.

Solution to λ^2 is an eigenvalue of A^2

Given $Av = \lambda v$, then:

$$\begin{aligned} A^2 v &= A(Av) \\ &= A(\lambda v) \\ &= \lambda(Av) \\ &= \lambda(\lambda v) \\ &= \lambda^2 v \end{aligned}$$

Thus, λ^2 is an eigenvalue of A^2 with the same eigenvector v.

Solution to λ^{-1} is an eigenvalue of A^{-1}

Assuming A is invertible, then:

$$\begin{aligned} A^{-1}(Av) &= A^{-1}(\lambda v) \\ (A^{-1}A)v &= \lambda A^{-1}v \\ Iv &= \lambda A^{-1}v \\ v &= \lambda A^{-1}v \\ A^{-1}v &= \lambda^{-1}v \end{aligned}$$

Hence, λ^{-1} is an eigenvalue of A^{-1}.

Solution to $\lambda + 1$ is an eigenvalue of $A + I$

Consider the identity matrix I:

$$\begin{aligned}(A+I)v &= Av + Iv \\ &= \lambda v + 1 \cdot v \\ &= (\lambda + 1)v\end{aligned}$$

Therefore, $\lambda + 1$ is an eigenvalue of $A + I$.

Eigenvectors of 2x2 matrices

Eigenvectors of 2-by-2 matrices

Consider a 2×2 matrix A with two distinct eigenvalues λ_1 and λ_2.

Explain why the columns of $A - \lambda_1 I$ are multiples of an eigenvector $\mathbf{x}_2$ associated with λ_2.

Eigenvectors and Eigenvalues

- For a matrix A with distinct eigenvalues λ_1 and λ_2, subtracting $\lambda_1 I$ from A zeroes out the component of the transformation in the direction of the eigenvector associated with λ_1.
- This leaves only the component that scales in the direction of the second eigenvector, hence why the columns are multiples of the second eigenvector.

Eigenvectors of 2×2 Matrices

We need to understand two key aspects:

1. Both columns of $A - \lambda_1 I$ being multiples of x_2 implies $A - \lambda_1 I$ transforms any vector into a vector in the direction of x_2.
2. The second eigenvector x_2 remains unchanged by the transformation $A - \lambda_1 I$ except for scaling by $\lambda_2 - \lambda_1$.

Detailed Solution

Let x_1 and x_2 be eigenvectors of A corresponding to λ_1 and λ_2 respectively. Then:

$$\begin{aligned} Ax_2 &= \lambda_2 x_2 \\ (A - \lambda_1 I)x_2 &= Ax_2 - \lambda_1 x_2 \\ &= \lambda_2 x_2 - \lambda_1 x_2 \\ &= (\lambda_2 - \lambda_1)x_2 \end{aligned}$$

Hence, the columns of $A - \lambda_1 I$ are scaled versions of x_2, which shows they are multiples of the eigenvector x_2.

Extending an Orthogonal Matrix: Problem Statement

New Q column

If Q is an $m \times n$ orthogonal matrix, then the projection matrix P for the column space of Q is $P = QQ^T$.

Suppose we add a new column $\mathbf{a}$ to Q to create a matrix $[Q \quad \mathbf{a}]$.

If we want to transform A into an orthogonal matrix $Q' = [Q \quad \mathbf{q}]$ via Gram-Schmidt, what is the vector $\mathbf{q}$?.

Gram-Schmidt Process

- The Gram-Schmidt process is used to orthogonalize a set of vectors in an inner product space.
- Starting with a set of linearly independent vectors, the process generates an orthogonal set that spans the same subspace.
- When we add a new vector to an orthogonal matrix, we want to ensure the new set remains orthogonal.

Gram Schmidt Process

Gram Schmidt orthonormalization process

Gram Schmidt Process

$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\
\mathbf{u}_4 &= \mathbf{v}_4 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_3}(\mathbf{v}_4), & \mathbf{e}_4 &= \frac{\mathbf{u}_4}{\|\mathbf{u}_4\|} \\
&\ \ \vdots & &\ \ \vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k), & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.
\end{aligned}
$$

Approach

To make q' orthogonal to the columns of Q, we must ensure it is perpendicular to the subspace spanned by Q.

- Project a onto the column space of Q to find the component of a in that space: $P = QQ^T a$.
- Subtract this projection from a to get a vector orthogonal to all columns of Q: $e = a - QQ^T a$.
- Normalize e to get the unit vector q': $q' = \frac{e}{||e||}$.

Detailed Solution

To find q', follow these steps:

1. Compute the projection matrix $P = QQ^T$.
2. Calculate the projection of a onto the space of Q: $P_a = QQ^T a$.
3. Subtract P_a from a to get the orthogonal vector e: $e = a - P_a$.
4. Normalize e to obtain q': $q' = \frac{e}{\|e\|}$.

This q' can now be appended to Q to form an orthogonal matrix Q'.

Subspaces and Linear Dependence

Subspaces and linear dependence

Let A be an $n \times n$ matrix of rank r and consider an eigenvector $\mathbf{x}$ with eigenvalue λ—that is, $A\mathbf{x} = \lambda\mathbf{x}$.

If $\lambda = 0$, then $\mathbf{x}$ is in the null space of A. If $\lambda \neq 0$, then $\mathbf{x}$ is in the column space of A.

The column space of A has dimension r. The null space of A has dimension $n - r$.

But not every $n \times n$ matrix will have n linearly independent eigenvectors. Why not?

Solution: Linear Independence of Eigenvectors

Consider the matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$. The vector $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$ lies in both the column space and the null space of A.

- The column space and null space can share vectors in square matrices
- In such matrices, there may not be enough linearly independent eigenvectors.
- The dimension of the null space is 1, but the multiplicity of the eigenvalue 0 is 2.
- This results in fewer than n independent eigenvectors for the matrix A.

Note: The algebraic multiplicity of an eigenvalue can exceed the dimension of the associated eigenspace, which is why a matrix may not have a full set of n linearly independent eigenvectors.

Least Square Approximations with orthogonality

Least Square Approximations with orthogonality

Suppose we want to construct a line $b = C + Dt$ that goes through

1. $b = 11$ at $t = -4$,
2. $b = 7$ at $t = -2$,
3. $b = 4$ at $t = 0$.
4. $b = -2$ at $t = 3$.

Construct A to set up the least squares equation.

$A =$

Compute the QR factorization for A.

$Q =$

$R =$

Find the least square solutions and fill in the value of C and D

$C =$ number (2 digits after decimal)

$D =$ number (2 digits after decimal)

Setting Up the Problem

We are given four points and want to find the best fit line $b = C + Dt$.

- The points are $(-4, 11)$, $(-2, 7)$, $(0, 4)$, and $(3, -2)$.
- We set up matrix A using t values and a vector b using b values.
- The matrix A has two columns: one for the intercept (all ones) and one for the slope (the t values).

Matrix A and vector b are:

$$A = \begin{bmatrix} 1 & -4 \\ 1 & -2 \\ 1 & 0 \\ 1 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} 11 \\ 7 \\ 4 \\ -2 \end{bmatrix}$$

Computing the QR Factorization

The QR factorization of matrix A decomposes it into:

- An orthogonal matrix Q.
- An upper triangular matrix R.

This factorization is used to solve the least squares problem more efficiently and with better numerical stability than directly solving $A^T Ax = A^T b$.

Solving the Least Squares System

With QR factorization $A = QR$, we solve the system:

- $Rx = Q^T b$ for vector x, which contains C and D.
- You get Q using Gram Schmidt
- We use the fact that Q is orthogonal, meaning $Q^T Q = I$.
- $A = QR$ which means $Q^T A = Q^T QR$

The solution x gives us the least squares estimates for C and D.

Results and Interpretation

After performing the QR factorization and solving the system, we get:

$$C = 3.63, \quad D = -1.83$$

Line of best fit:

$$b = 3.63 - 1.83t$$

This line minimizes the sum of the squares of the vertical distances (residuals) between the given points and the line.