Matrix Decompositions

April 13, 2023

1 Eigendecomposition and Diagonalization

A diagonal matrix is a matrix where all elements except the diagonal ones are zero.

$$\boldsymbol{D} = \begin{bmatrix} c_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & c_n \end{bmatrix} \tag{1}$$

These matrices allow fast computation of determinants, powers and inverses.

Two matrices $\boldsymbol{A}$ and $\boldsymbol{D}$ are similar if there exists an invertible matrix $\boldsymbol{P}$, such that $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. More specifically, we look at matrices $\boldsymbol{A}$ that are similar to diagonal matrices $\boldsymbol{D}$ that contain the eigenvalues of $\boldsymbol{A}$ on the diagonal.

Definition - Diagonalizable
A matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is diagonalizable if it is similar to a diagonal matrix i.e., if there exists an invertible matrix $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ such that $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, let $\lambda_1, \dots, \lambda_n$ be a set of scalars, and let $\boldsymbol{p}_1, \dots, \boldsymbol{b}_n$ be a set of vectors in $\mathbb{R}^n$. We define $\boldsymbol{P} := [\boldsymbol{p}_1, \dots, \boldsymbol{p}_n]$ and let $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ be a diagonal matrix with diagonal elements $\lambda_1, \dots, \lambda_n$. Then we can show that

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{D} \tag{2}$$

if and only id $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{p}_1, \dots, \boldsymbol{p}_n$ are the eigenvectors of $\boldsymbol{A}$.
We can see that statements holds that

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{A}[\boldsymbol{p}_1, \dots, \boldsymbol{p}_2] = [\boldsymbol{A}\boldsymbol{p}_1, \dots, \boldsymbol{A}\boldsymbol{p}_n] \tag{3}$$

$$\boldsymbol{P}\boldsymbol{D} = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_n] \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix} = [\lambda_1\boldsymbol{p}_1, \dots, \lambda_n\boldsymbol{p}_n] \tag{4}$$

Thus it implies that

$$\begin{gathered} \boldsymbol{A}\boldsymbol{p}_1 = \lambda_1\boldsymbol{p}_1 \\ \vdots \\ \boldsymbol{A}\boldsymbol{p}_n = \lambda_n\boldsymbol{p}_n \end{gathered} \tag{5}$$

Therefore the columns of $\boldsymbol{P}$ must be the eigenvectors of $\boldsymbol{A}$.

Theorem - Eigendecomposition
A square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ can be factored into

$$\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} \tag{6}$$

where $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ and $\boldsymbol{D}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\boldsymbol{A}$, if and only if the eigenvectors of $\boldsymbol{A}$ form the basis of $\mathbb{R}^n$.
This theorem implies that only non-defective matrices can be diagonlized and that the columns of $\boldsymbol{P}$ are the n eigenvectors of $\boldsymbol{A}$. For symmetric matrices we get obtain even stronger outcomes for the eigenvalue decomposition.

Theorem
A symmetric matrix $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ can always be diagonlized.
This theorem implies that the matrix $\boldsymbol{P}$ is orthgonal such that $\boldsymbol{D} = \boldsymbol{P}^T \boldsymbol{A} \boldsymbol{P}$.

Useful properties:

- Diagonal matrices $\boldsymbol{D}$ can efficiently be raised to a power. Therefore, we can find a power for a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ via the eigenvalue decomposition (if it exists) so that

$$\boldsymbol{A}^k = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^k = \boldsymbol{P}\boldsymbol{D}^k\boldsymbol{P}^{-1} \tag{7}$$

  Computing $\boldsymbol{D}^k$ is efficient because we apply this operation individually to any diagonal element.

- Assume that the eigendecomposition $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$ exists. Then,

$$det(\boldsymbol{A}) = det(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}) = det(\boldsymbol{P})det(\boldsymbol{D})det(\boldsymbol{P}^{-1}) = det(\boldsymbol{D}) = \prod_{i=1}^{n} d_{ii} \tag{8}$$

  allows for an efficient computation of the determinant of $\boldsymbol{A}$.