Spectral Theory for Real Symmetric Matrices

Gil Cohen

October 17, 2021

Overview

Eigenvalues

Recall that a nonzero vector ψ is an eigenvector of a matrix $\mathbf{M}$ with eigenvalue λ if

$$\mathbf{M}\psi = \lambda\psi.$$

Hence, λ is an eigenvalue of $\mathbf{M}$ if

- $\lambda\mathcal{I} - \mathbf{M}$ is singular;
- λ is a root of the characteristic polynomial of $\mathbf{M}$, $\det(x\mathcal{I} - \mathbf{M})$.

Quick important corollaries:

- $\mathbf{M}$ has n eigenvalues in $\mathbb{C}$, counted with multiplicities.
- The product of eigenvalues $\prod_i \lambda_i = \det \mathbf{M}$.

The Spectral Theorem

Theorem (The Spectral Theorem)

Let $\mathbf{M}$ be an $n \times n$ real, symmetric matrix. Then there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (not necessarily distinct) and n mutually orthogonal unit vectors $\psi_1, \ldots, \psi_n$ such that ψ_i is an eigenvector of $\mathbf{M}$ of eigenvalue λ_i.

Spectral Decomposition

Corollary (Spectral decomposition)

Let $\mathbf{M}$ be a real, symmetric $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding orthonormal eigenvectors $\psi_1, \ldots, \psi_n$. Then,

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T = \sum_{i=1}^{n} \lambda_i \psi_i \psi_i^T$$

where $\mathbf{U} = (\psi_1, \ldots, \psi_n)$ and $\mathbf{\Sigma} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Thinking of $\mathbf{M}$ as an operator, take $\mathbf{x} \in \mathbb{R}^n$ and write $\mathbf{x} = \sum_i c_i \psi_i$ where $\sum_i c_i^2 = \|\mathbf{x}\|_2^2$. We have that

$$\mathbf{M}\mathbf{x} = \sum_i c_i \mathbf{M}\psi_i = \sum_i \lambda_i c_i \psi_i.$$

The spectral decomposition is useful for taking powers

$$\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T = \sum_{i=1}^{n} \lambda_i \psi_i \psi_i^T$$

$$\mathbf{M}^2 = \mathbf{U}\boldsymbol{\Sigma}^2\mathbf{U}^T = \sum_{i=1}^{n} \lambda_i^2 \psi_i \psi_i^T$$

If $\lambda_i \neq 0$ for all i, then

$$\mathbf{M}^{-1} = \mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{U}^T = \sum_{i=1}^{n} \frac{1}{\lambda_i} \psi_i \psi_i^T.$$

If $\mathbf{M}$ is singular, we can still define the pseudo-inverse (aka the Moore–Penrose inverse) by

$$\mathbf{M}^\dagger = \sum_{i:\lambda_i \neq 0} \frac{1}{\lambda_i} \psi_i \psi_i^T.$$

Positive (semi)definite matrices

Definition

A real symmetric matrix $\mathbf{M}$ is positive semidefinite (PSD) if all its eigenvalues are non-negative. It is positive definite (PD) if its eigenvalues are strictly positive.

For a PSD $\mathbf{M}$,

$$\mathbf{M} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^T = \sum_{i=1}^{n} \lambda_i \psi_i \psi_i^T$$

$$\sqrt{\mathbf{M}} = \mathbf{U}\sqrt{\mathbf{\Sigma}}\mathbf{U}^T = \sum_{i=1}^{n} \sqrt{\lambda_i} \psi_i \psi_i^T$$

Overview

The trace

Theorem

The trace function is cyclic:

$$\mathrm{Tr}(\mathbf{MN}) = \mathrm{Tr}(\mathbf{NM}).$$

Theorem

Let $\mathbf{M}$ be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then,

$$\mathrm{Tr}(\mathbf{M}) = \sum_{i=1}^{n} \lambda_i.$$

The trace

Theorem

Let $\mathbf{M}$ be an $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then,

$$\mathrm{Tr}(\mathbf{M}) = \sum_{i=1}^{n} \lambda_i.$$

Proof.

$$\mathrm{Tr}(\mathbf{M}) = \mathrm{Tr}(\mathbf{U}\Sigma\mathbf{U}^T) = \mathrm{Tr}(\mathbf{U}^T\mathbf{U}\Sigma) = \mathrm{Tr}(\Sigma) = \sum_{i=1}^{n} \lambda_i.$$

□

Cyclicity and the characteristic polynomial

Lemma

For $n \times n$ matrices $\mathbf{A}, \mathbf{B}$,

$$\phi_{\mathbf{AB}}(x) = \phi_{\mathbf{BA}}(x).$$

Hence, the spectrum of $\mathbf{AB}$ is the spectrum of $\mathbf{BA}$.

More generally, if $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ an $m \times n$ matrix with $n > m$ then

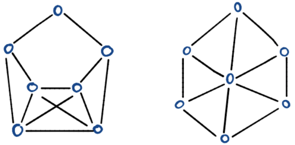
$$\phi_{\mathbf{AB}}(x) = x^{n-m}\phi_{\mathbf{BA}}(x).$$

In particular, the spectrum remains the same up to the change in kernel.

Overview

Cospectral graphs

Graphs G, H with the same sequence of eigenvalues of their respective $\mathbf{M}_G, \mathbf{M}_H$ are called cospectral. Isomorphic graphs are cospectral but not the other way around.



The adjacency matrices of both graphs have the same characteristic polynomial

$$(x+2)(x+1)^2(x-1)^2(x^2-2x-6)$$

Spectral properties of a graph

We say that a property of a graph is a spectral property if it is determined by its eigenvalues (its spectrum).

Say G is a graph with e edges. As $\mathrm{Tr}(\mathbf{M}_G) = \sum_i \lambda_i$,

$$\sum_i \lambda_i^2 = \mathrm{Tr}(\mathbf{M}_G^2) = \sum_v \deg v = 2e.$$

Hence, the number of edges is a spectral property.

Overview

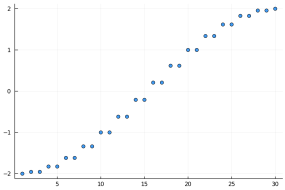
The spectrum of the cycle graph

Lemma

Let G be the cycle graph on $V = [n]$. Let $\omega \in \mathbb{C}$ be an n^{th} root of unity. Then, for every $i \in [n]$,

$$\mu_i = \omega^i + \omega^{-i}$$

is an eigenvalue of $\mathbf{M}_G$ with eigenvector ψ_i with j^{th} entry $(\psi_i)_j = \omega^{ij}$.



Overview

The Rayleigh quotient

Definition (The Rayleigh quotient)

The Rayleigh quotient of a vector $\mathbf{x}$ with respect to a matrix $\mathbf{M}$ is defined by

$$\frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Question

What is the Rayleigh quotient of an eigenvector of $\mathbf{M}$?

Question

What is the largest value that the Rayleigh quotient can attain?

The Rayleigh quotient

Note that if μ_2 is the second largest eigenvalue of $\mathbf{M}$ and ψ_1 is the eigenvector corresponding to the largest eigenvalue μ_1 then

$$\mu_2 = \max_{0 \neq \mathbf{x} \perp \psi_1} \frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Even without referring to ψ_1 we can write

$$\mu_2 = \max_{\substack{S \subseteq \mathbb{R}^n \\ \dim S = 2}} \min_{\substack{\mathbf{x} \in S \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$$

The Courant-Fischer Theorem

Theorem (The Courant-Fischer Theorem)

Let $\mathbf{M}$ be a symmetric matrix with eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. Then,

$$\mu_k = \max_{\substack{S \subseteq \mathbb{R}^n \\ \dim S = k}} \min_{\substack{\mathbf{x} \in S \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$$

$$= \min_{\substack{T \subseteq \mathbb{R}^n \\ \dim T = n-k+1}} \max_{\substack{\mathbf{x} \in T \\ \mathbf{x} \neq 0}} \frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Overview

The Perron-Frobenius Theorem (for symmetric matrices)

Theorem

Let G be a connected undirected graph with adjacency matrix $\mathbf{M}$ and corresponding eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. Then,

- μ_1 has a strictly positive eigenvector.
- $\mu_1 \geq -\mu_n$. Equality holds if and only if G is bipartite.
- $\mu_1 > \mu_2$.

Summary