

Linear algebra background

The material in this handout is taken from: Luca Trevisan, Lecture Notes on Graph Partitioning, Expanders and Spectral Methods, 2016.

Given a real $n \times n$ matrix M , if $M\mathbf{u} = \lambda\mathbf{u}$ for some $\lambda \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, then \mathbf{u} is an eigenvector of M with eigenvalue λ (we also say that \mathbf{u} is an eigenvector of λ). Note that eigenvectors can be rescaled without changing the equation $M\mathbf{u} = \lambda\mathbf{u}$, hence we conventionally assume they have unit length.

Note that λ is an eigenvalue for M if and only if there exists $\mathbf{x} \neq \mathbf{0}$ such that $(M - \lambda I)\mathbf{x} = \mathbf{0}$, where I is the $n \times n$ identity matrix. The equation $(M - \lambda I)\mathbf{x} = \mathbf{0}$ holds for $\mathbf{x} \neq \mathbf{0}$ if and only if $M - \lambda I$ is singular, which is equivalent to $\det(M - \lambda I) = 0$. Since $\det(M - \lambda I)$ is a n -th degree univariate polynomial in λ , it has exactly n solutions by the fundamental theorem of algebra. This shows that every square matrix has n eigenvalues (not all necessarily distinct). Some of these eigenvalues, however, may correspond to solutions of $\det(M - \lambda I) = 0$ in the complex plane. The next result guarantees that at least one eigenvalue is real when M is symmetric.

Fact 1 (proof omitted) *If M is symmetric, then there exists $\lambda \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that $M\mathbf{u} = \lambda\mathbf{u}$.*

Fact 2 *If M is symmetric then any two eigenvectors corresponding to distinct eigenvalues are orthogonal.*

PROOF. Let \mathbf{x} be an eigenvector of λ and \mathbf{y} an eigenvector of λ' with $\lambda \neq \lambda'$. Since M is symmetric, $(M\mathbf{x})^\top \mathbf{y} = \mathbf{x}^\top M\mathbf{y}$. On the other hand, $(M\mathbf{x})^\top \mathbf{y} = \lambda \mathbf{x}^\top \mathbf{y}$ and $\mathbf{x}^\top M\mathbf{y} = \lambda' \mathbf{x}^\top \mathbf{y}$. Since $\lambda \neq \lambda'$, it must be that $\mathbf{x}^\top \mathbf{y} = 0$, which means that \mathbf{x} and \mathbf{y} are orthogonal. \square

Theorem 3 (Spectral Theorem) *Let $M \in \mathbb{R}^{n \times n}$ be a real symmetric matrix. Then there exists n (not necessarily distinct) real numbers $\lambda_1, \dots, \lambda_n$ and n orthonormal real vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ such that \mathbf{u}_i is an eigenvector of λ_i .*

PROOF. The proof is by induction on n . If $n = 1$, then M is a scalar. Hence, any nonzero $x \in \mathbb{R}$ is an eigenvector of M with eigenvalue M because $Mx = Mx$.

Assume now that the statement holds for $n - 1$. By Fact 1, there exist an eigenvalue $\lambda_n \in \mathbb{R}$ with eigenvector $\mathbf{x}_n \in \mathbb{R}^n$.

Claim. \mathbf{y} orthogonal to \mathbf{x}_n implies $M\mathbf{y}$ is orthogonal to \mathbf{x}_n .

Indeed, $\mathbf{x}_n^\top M\mathbf{y} = (M\mathbf{x}_n)^\top \mathbf{y} = \lambda_n \mathbf{x}_n^\top \mathbf{y} = 0$.

Now let V be the $(n-1)$ -dimensional subspace of \mathbb{R}^n that contains all the vectors orthogonal to \mathbf{x}_n . Now choose an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ for V and let $B = [\mathbf{u}_1, \dots, \mathbf{u}_{n-1}]$. By construction, B maps \mathbb{R}^{n-1} to V and BB^\top maps \mathbb{R}^n to V . In particular, $BB^\top \mathbf{z} = \mathbf{z}$ for all $\mathbf{z} \in V$. We now apply the inductive hypothesis to the $(n-1) \times (n-1)$ symmetric matrix $M' = B^\top MB$ and find real eigenvalues $\lambda_1, \dots, \lambda_{n-1}$ and orthonormal eigenvectors $\mathbf{y}_1, \dots, \mathbf{y}_{n-1} \in \mathbb{R}^{n-1}$. For $i = 1, \dots, n-1$ we have $M'\mathbf{y}_i = B^\top MB\mathbf{y}_i = \lambda_i \mathbf{y}_i$. Therefore, $BB^\top MB\mathbf{y}_i = \lambda_i B\mathbf{y}_i$. Since $\mathbf{y}_i \in \mathbb{R}^{n-1}$ and B maps \mathbb{R}^{n-1} to V , $B\mathbf{y}_i$ is orthogonal to \mathbf{x}_n and, by the above claim, $MB\mathbf{y}_i$ is orthogonal to \mathbf{x}_n . Therefore $\lambda_i B\mathbf{y}_i = BB^\top MB\mathbf{y}_i = MB\mathbf{y}_i$. If we now define $\mathbf{x}_i = B\mathbf{y}_i$ for $i = 1, \dots, n-1$, then we have $M\mathbf{x}_i = \lambda_i \mathbf{x}_i$. To finish up, note that by construction \mathbf{x}_n is orthogonal to $\mathbf{x}_1, \dots, \mathbf{x}_{n-1}$. Moreover, for any $1 \leq i < j \leq n-1$, $\mathbf{x}_i^\top \mathbf{x}_j = (B\mathbf{y}_i)^\top (B\mathbf{y}_j) = \mathbf{y}_i^\top B^\top B\mathbf{y}_j = \mathbf{y}_i^\top \mathbf{y}_j = 0$. Hence we have found n eigenvalues with n eigenvectors. \square

Corollary 4 Let $M \in \mathbb{R}^{n \times n}$ be a real symmetric matrix. Then

$$M = U\Lambda U^\top = \sum_{i=1}^n \lambda_i \mathbf{u}_i \mathbf{u}_i^\top$$

where $U = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Here $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are the real eigenvalues of M and $\mathbf{u}_1, \dots, \mathbf{u}_n \in \mathbb{R}^n$ are the corresponding eigenvectors.

PROOF. Note that $MU = [\lambda_1 \mathbf{u}_1, \dots, \lambda_n \mathbf{u}_n]$ because $M\mathbf{u}_i = \lambda_i \mathbf{u}_i$ for each $i = 1, \dots, n$. Hence $MU = U\Lambda$ where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Since U is orthonormal, $U^{-1} = U^\top$ and $UU^\top = I$. Therefore $M = MUU^\top = U\Lambda U^\top$. \square

Theorem 5 (Variational characterization of eigenvalues — proof omitted) Let $M \in \mathbb{R}^{n \times n}$ be a real symmetric matrix, and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be its real eigenvalues. For $k < n$ let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be orthonormal vectors such that $M\mathbf{u}_i = \lambda_i \mathbf{u}_i$ for $i = 1, \dots, k$. Then

$$\lambda_{k+1} = \min_{\substack{\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \\ \mathbf{u} \perp \{\mathbf{u}_1, \dots, \mathbf{u}_k\}}} \frac{\mathbf{u}^\top M \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}$$

and any minimizer \mathbf{u} is an eigenvector of λ_{k+1} .

The ratio in the right-hand side of the above equation is called **Rayleigh quotient**. Note that, in particular,

$$\lambda_1 = \min_{\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^\top M \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}.$$

Also, because $-M$ has eigenvalues $-\lambda_n \leq -\lambda_{n-1} \leq \dots \leq -\lambda_1$,

$$-\lambda_n = \min_{\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^\top (-M) \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = - \max_{\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^\top M \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}$$

and therefore

$$\lambda_n = \max_{\mathbf{u} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^\top M \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}$$

A symmetric matrix M is **positive semidefinite** if $\mathbf{x}^\top M \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

Fact 6 *The eigenvalues of a positive semidefinite matrix are all nonnegative.*

PROOF. As the denominator of the Rayleigh quotient is clearly always positive, Theorem 5 implies that the sign of each eigenvalue is determined by the sign of $\mathbf{x}^\top M \mathbf{x}$. \square

We conclude with a different but equally important characterization of eigenvalues.

Theorem 7 (Courant-Fischer — proof omitted) *Let M be a real symmetric matrix with real eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Then*

$$\lambda_k = \min_{S: \dim(S)=k} \max_{\mathbf{u} \in S \setminus \{\mathbf{0}\}} \frac{\mathbf{u}^\top M \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} \quad k = 1, \dots, n$$

where the minimum is over all subspaces S of dimension k .