Symmetric matrices

Properties of real symmetric matrices

- Recall that a matrix $A \in \mathbb{R}^{n \times n}$ is **symmetric** if $A^T = A$.
- For real symmetric matrices we have the following two crucial properties:
- All eigenvalues of a real symmetric matrix are real.
- Eigenvectors corresponding to distinct eigenvalues are orthogonal.
- To show these two properties, we need to consider complex matrices of type $A \in \mathbb{C}^{n \times n}$, where $\mathbb{C}$ is the set of complex numbers $z = x + iy$ where x and y are the real and imaginary part of z and $i = \sqrt{-1}$.
- $\mathbb{C}^n$ is the set of n-column vectors with components in $\mathbb{C}$ and similarly $\mathbb{C}^{n \times n}$ is the set of $n \times n$ matrices with complex numbers as its entries.
- We write the complex conjugate of z as $z^* = x - iy$. For $u \in \mathbb{C}^n$ and $A \in \mathbb{C}^{n \times n}$, we denote by $u^* \in \mathbb{C}^n$ and $A^* \in \mathbb{C}^{n \times n}$, their complex conjugates, obtained by taking the complex conjugate of each of their components.

Properties of real symmetric matrices

- We write the complex conjugate of z as $z^* = x - iy$.
- For $u \in \mathbb{C}^n$, we denote by $u^* \in \mathbb{C}^n$ its complex conjugate, obtained by taking the complex conjugate of each of its components, i.e., $(u^*)_i = (u_i)^*$.
- Similarly, for $A \in \mathbb{C}^{n \times n}$, we denote by $A^* \in \mathbb{C}^{n \times n}$, the complex conjugate of A, obtained by taking the complex conjugate of each of its entries, i.e., $(A^*)_{ij} = (A_{ij})^*$.
- Note that for complex numbers we have $z_1 = z_2$ iff $z_1^* = z_2^*$.
- This property clearly extends to complex vectors and matrices:
- For $u, v \in \mathbb{C}^n$ we have $u = v$ iff $u^* = v^*$ and for $A, B \in \mathbb{C}^{n \times n}$, we have $A = B$ iff $A^* = B^*$.
- Furthermore, $(Au)^* = A^* u^*$ and $(A^*)^T = (A^T)^*$.

Eigenvalues of a symmetric real matrix are real

- Let $\lambda \in \mathbb{C}$ be an eigenvalue of a symmetric $A \in \mathbb{R}^{n\times n}$ and let $u \in \mathbb{C}^n$ be a corresponding eigenvector:

$$Au = \lambda u. \tag{1}$$

- Taking complex conjugates of both sides of (1), we obtain:

$$A^* u^* = \lambda^* u^*, \text{ i.e., } Au^* = \lambda^* u^*. \tag{2}$$

- Now, we pre-multiply (1) with $(u^*)^T$ to obtain:

$$\begin{aligned}
\lambda(u^*)^T u &= (u^*)^T(Au) = ((u^*)^T A)u & \\
&= (A^T u^*)^T u & \text{since } (Bv)^T = v^T B^T \\
&= (Au^*)^T u & \text{since } A^T = A \\
&= (\lambda^* u^*)^T u = \lambda^*(u^*)^T u. & \text{using (2)}
\end{aligned}$$

- Thus, $(\lambda - \lambda^*)(u^*)^T u = 0$.
- But u, being an eigenvector is non-zero and $(u^*)^T u = \sum_{i=1}^n u_i^* u_i > 0$ since at least one of the components of u is non-zero and for any complex number $z = a + ib$ we have $z^* z = a^2 + b^2 \geq 0$.
- Hence $\lambda = \lambda^*$, i.e., λ and hence u are both real.

Eigenvectors of distinct eigenvalues of a symmetric real matrix are orthogonal

- Let A be a real symmetric matrix.
- Let $Au_1 = \lambda_1 u_1$ and $Au_2 = \lambda_2 u_2$ with u_1 and u_2 non-zero vectors in $\mathbb{R}^n$ and $\lambda_1, \lambda_2 \in \mathbb{R}$.
- Pre-multiplying both sides of the first equation above with u_2^T, we get:

$$\begin{aligned} \lambda u_2^T u_1 = u_2^T (Au_1) = (u_2^T A) u_1 &= (A^T u_2)^T u_1 \\ &= (Au_2)^T u_1 = \lambda_2 u_2^T u_1. \end{aligned}$$

- Thus, $(\lambda_1 - \lambda_2) u_2^T u_1 = 0$.
- Therefore, $\lambda_1 \neq \lambda_2$ implies: $u_2^T u_1 = 0$ as required.
- If an eigenvalue λ has multiplicity m say then we can always find a set of m orthonormal eigenvectors for λ.
- We conclude that by normalizing the eigenvectors of A, we get an orthonormal set of vectors $u_1, u_2, \ldots, u_n$.

Properties of positive definite symmetric matrices

- Suppose $A \in \mathbb{R}^n$ is a symmetric positive definite matrix, i.e., $A = A^T$ and

$$\forall x \in \mathbb{R}^n \setminus \{0\}.\, x^T Ax > 0. \tag{3}$$

- Then we can easily show the following properties of A.
- All diagonal elements are positive: In (3), put x with $x_j = 1$ for $j = i$ and $x_j = 0$ for $j \neq i$, to get $A_{ii} > 0$.
- The largest element in magnitude in the entire matrix occurs in the diagonal: Fix $i \neq j$ between 1 and n. In (3), put x with $x_k = 1$ for $k = i$, $x_k = \pm 1$ for $k = j$ and $x_k = 0$ for $j \neq k \neq i$, to get $|A_{ij}| < \max(A_{ii}, A_{jj})$.
- All leading principle minors (i.e., the 1×1, 2×2, 3×3, $\ldots, m \times m$ matrices in the upper left corner) are positive definite: In (3), put x with $x_k = 0$ for $k > m$ to prove that the top left $m \times m$ matrix is positive definite.

Spectral decomposition

- We have seen in the previous pages and in lecture notes that if $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix then it has an orthonormal set of eigenvectors $u_1, u_2, \dots, u_n$ corresponding to (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, then we have:
- The **spectral decomposition**: $Q^T A Q = \Lambda$ where
- $Q = [u_1, u_2, \dots, u_n]$ is an orthogonal matrix with $Q^{-1} = Q^T$ and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is diagonal.
- Assume $A \in \mathbb{R}^n$ reprsents the linear map $f : \mathbb{R}^n \to \mathbb{R}^n$ in the standard basis of $\mathbb{R}^n$.
- Then, the matrix $S := Q^{-1} \in \mathbb{R}^{n \times n}$ is the matrix for the change of basis into one in which f is reprsented by $B := \Lambda$. More generally, $B = SAS^{-1}$:

$$
\begin{array}{ccc}
\text{old coordinates} & \xrightarrow{A} & \text{old coordinates} \\
S \big\downarrow & & \big\downarrow S \\
\text{new coordinates} & \xrightarrow{B} & \text{new coordinates}
\end{array}
$$

Singular value decomposition (SVD) I

- Let $A \in \mathbb{R}^{m \times n}$ be an arbitrary matrix.
- Then $A^T A \in \mathbb{R}^{n \times n}$ and $AA^T \in \mathbb{R}^{m \times m}$ are symmetric matrices.
- They are also positive semi-definite since for example $x^T A^T A x = (Ax)^T (Ax) = (\|Ax\|_2)^2 \geq 0$.
- We will show that $A = USV^T$, called the SVD of A, where $V \in \mathbb{R}^{n \times n}$ and $U \in \mathbb{R}^{m \times m}$ are orthogonal matrices whereas the matrix $S = U^T A V \in \mathbb{R}^{m \times n}$ is diagonal with $S = \text{diag}(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_p)$ where $p = \min(m, n)$ and the non-negative numbers $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq \sigma_p \geq 0$ are the **singular values** of A.
- If r is the rank of A then A has exactly r positive singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq \sigma_r > 0$ with $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_p = 0$.

Singular value decomposition II

- Note that if the SVD for A as above exists then, since $U^T U = I_m$, we have $A^T A = VS^T U^T USV^T = VS^T SV^T$, where $S^T S = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_p^2) \in \mathbb{R}^{n \times n}$ is a diagonal matrix, thus giving the spectral decomposition of the positive semi-definite matrix $A^T A$.
- This gives us a method to find the SVD of A.
- Obtain the eigenvalues $\sigma_1^2 \geq \sigma_2^2 \geq \sigma_3^2 \geq \ldots \geq \sigma_p^2 \geq 0$ and the corresponding eigenvectors $v_1, \cdots, v_p$ of $A^T A$. If $p < n$, the other eigenvalues of A are zero with corresponding eigenvectors $v_{p+1}, \cdots, v_n$ which make the orthogonal matrix $V = [v_1, \cdots, v_n]$.
- From the SVD we have $AV = US$, thus when $\sigma_i > 0$, i.e., for $1 \leq i \leq r$, we get $\frac{1}{\sigma_i} Av_i = u_i$.
- Extend the set $u_1, \cdots, u_r$ to an orthonormal basis $u_1, \cdots, u_r, \cdots u_m$ of $\mathbb{R}^m$ which gives the orthogonal matrix $U = [u_1, \cdots, u_m]$.

Singular value decomposition III

- Note the following useful facts.
- For $1 \leq i \leq r$, the vector u_i is an eigenvector of AA^T with eigenvalue σ_i^2. Check!
- AA^T is similar to SS^T (with identical eigenvalues) and A^TA is similar to S^TS (with identical eigenvalues).
- The diagonal elements of the diagonal matrices $S^TS \in \mathbb{R}^{n \times n}$ and $SS^T \in \mathbb{R}^{m \times m}$ are $\sigma_1^2, \sigma_2^2, \sigma_3^2, \ldots, \sigma_p^2$ followed by $n - p$ zeros and $m - p$ zeros respectively.
- The singular values of A are the positive square roots of the eigenvalues of AA^T or A^TA.