Mathematics for Computer Science
Linear Algebra

Lecture 18: Orthogonal matrices and spectral decomposition

Andrei Krokhin

March 4, 2021

Contents for today's lecture

- Orthogonal matrices: definition, properties, and characterisations
- Orthogonal diagonalisation: a characterisation and an algorithm
- Spectral decomposition of symmetric matrices

Reminder from earlier lectures

- Orthonormal = consisting of unit vectors, which are pairwise orthogonal
- Every finite-dimensional inner product space V has an orthonormal basis, which can be constructed from any basis of V via the Gram-Schmidt process.
- For any column vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ with Euclidean inner product, we have

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T \mathbf{v}.$$

Orthogonal matrices: definition

Definition

A square matrix Q is called orthogonal if $Q^T = Q^{-1}$ (equivalently, $Q^T Q = I$).

Example: Rotation and reflection matrices in $\mathbb{R}^2$ are orthogonal. Easy to check:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Example: All permutation matrices are orthogonal.

Rows and columns in orthogonal matrices

Theorem

For any $n \times n$ matrix Q, TFAE:

1. *Q is orthogonal.*
2. *The rows of Q form an orthonormal basis in $\mathbb{R}^n$ (with the Euclidean inner product).*
3. *The columns of Q form an orthonormal basis in $\mathbb{R}^n$ (with the Euclidean inner product).*

Proof.

We prove (1) $\Leftrightarrow$ (3), the proof of (1) $\Leftrightarrow$ (2) is similar.

If $Q = [\mathbf{u}_1 | \dots | \mathbf{u}_n]$ then each entry (i, j) in the product $Q^T Q$ is $\langle \mathbf{u}_i, \mathbf{u}_j \rangle$.

Hence, $Q^T Q = I$ iff we have $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $i \neq j$ and $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 1$ for all $i = j$, which holds iff the set $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is orthonormal.

Any orthonormal set with n vectors in $\mathbb{R}^n$ is a basis. $\square$

Properties of orthogonal matrices

Theorem

1. *The transpose of an orthogonal matrix is also orthogonal*
2. *The inverse of an orthogonal matrix is also orthogonal*
3. *A product of orthogonal matrices is also orthogonal.*
4. *If Q is orthogonal then $det(Q) = 1$ or $det(Q) = -1$.*

All four proofs are very easy (one-line) exercises.

Orthogonal matrices as linear operators

Theorem

For any $n \times n$ matrix Q, TFAE:

1. *Q is orthogonal.*
2. $||Q\mathbf{x}|| = ||\mathbf{x}||$ *for all* $\mathbf{x} \in \mathbb{R}^n$.
3. $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ *for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Proof.

(1) $\Rightarrow$ (3). If Q is orthogonal, $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = (Q\mathbf{x})^T Q\mathbf{y} = \mathbf{x}^T Q^T Q\mathbf{y} = \mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y}\rangle$.
(3) $\Rightarrow$ (2). Obvious, use (3) with $\mathbf{x} = \mathbf{y}$.
(2) $\Rightarrow$ (1). Note: if $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis then $Q = [Q\mathbf{e}_1|\dots|Q\mathbf{e}_n]$.
We have $||Q\mathbf{e}_i|| = ||\mathbf{e}_i|| = 1$ for all i, and if $i \neq j$ then $\langle Q\mathbf{e}_i, Q\mathbf{e}_j\rangle = 0$ because

$$2 = ||\mathbf{e}_i + \mathbf{e}_j||^2 = ||Q(\mathbf{e}_i + \mathbf{e}_j)||^2 = ||Q\mathbf{e}_i + Q\mathbf{e}_j||^2 = \langle Q\mathbf{e}_i + Q\mathbf{e}_j, Q\mathbf{e}_i + Q\mathbf{e}_j\rangle =$$
$$||Q\mathbf{e}_i||^2 + 2\langle Q\mathbf{e}_i, Q\mathbf{e}_j\rangle + ||Q\mathbf{e}_j||^2 = ||\mathbf{e}_i||^2 + 2\langle Q\mathbf{e}_i, Q\mathbf{e}_j\rangle + ||\mathbf{e}_j||^2 = 2 + 2\langle Q\mathbf{e}_i, Q\mathbf{e}_j\rangle.$$

The columns of Q form an orthonormal set, hence Q is orthogonal. $\square$

Orthogonal diagonalisation

Let A and B be $n \times n$ matrices. Recall:

- A and B are similar if there is an invertible matrix P such that $P^{-1}AP = B$.
- If B above can be chosen to be diagonal then we say that A is diagonalisable and that P diagonalises A.
- We proved: A is diagonalisable iff it has n linearly independent eigenvectors.

If, in addition, P can be chosen to be orthogonal then we say, respectively, that

- A and B are orthogonally similar.
- A is orthogonally diagonalisable and that P orthogonally diagonalises A (and then $P^T AP = B$, since $P^T = P^{-1}$).

Question: Which matrices are orthogonally diagonalisable?

The spectral theorem

Theorem (Spectral theorem)

For any $n \times n$ matrix A, TFAE:

1. *A is orthogonally diagonalisable,*
   i.e. $A = QDQ^T$ for some orthogonal matrix Q and some diagonal matrix D.
2. *A has an orthonormal set of n eigenvectors.*
3. *A is symmetric.*

Proof.

The proof of (1) $\Leftrightarrow$ (2) is the same as for the general diagonalisation (+ the fact that a matrix is orthogonal iff it has orthonormal columns).

(1) $\Rightarrow$ (3) If $A = QDQ^T$ then $A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$.

(3) $\Rightarrow$ (1) Proof is omitted. $\square$

Eigen-properties of symmetric matrices

Theorem

If A is a symmetric matrix then

1. *all (complex) eigenvalues A are real, and*
2. *eigenvectors from different eigenspaces are orthogonal.*

Proof.

We proved (1) in lecture 14 (about complex vector spaces).

To prove (2), assume that λ_1 and λ_2 are two different eigenvalues of A, and take any two eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ corresponding to λ_1 and λ_2, respectively. We have

$$\lambda_1\langle \mathbf{u}_1, \mathbf{u}_2\rangle = \langle \lambda_1\mathbf{u}_1, \mathbf{u}_2\rangle = \langle A\mathbf{u}_1, \mathbf{u}_2\rangle = (A\mathbf{u}_1)^T\mathbf{u}_2 =$$

$$\mathbf{u}_1^T A^T\mathbf{u}_2 = \mathbf{u}_1^T A\mathbf{u}_2 = \langle \mathbf{u}_1, A\mathbf{u}_2\rangle = \langle \mathbf{u}_1, \lambda_2\mathbf{u}_2\rangle = \lambda_2\langle \mathbf{u}_1, \mathbf{u}_2\rangle.$$

Since $(\lambda_1 - \lambda_2)\langle \mathbf{u}_1, \mathbf{u}_2\rangle = 0$ and $\lambda_1 \neq \lambda_2$, we get $\langle \mathbf{u}_1, \mathbf{u}_2\rangle = 0$, as required. □

How to orthogonally diagonalise a symmetric matrix

Algorithm (assuming A is symmetric):

Step 1. Find the eigenvalues and a basis in each eigenspace of A.

Step 2. Apply Gram-Schmidt to each basis to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix Q whose columns are the vectors found in Step 2. Q will orthogonally diagonalise A and the diagonal of $D = Q^T AQ$ contains the eigenvalues of A in the same order as eigenvectors in Q.

Remarks:

- Steps 1 and 3 were also part of the (general) diagonalisation algorithm.
- The columns of Q will form an orthonormal set because eigenvectors from different eigenspaces are orthogonal (and the rest is guaranteed by Step 2).
- This algorithm applies only to symmetric matrices and always returns a diagonalisation, unlike the general diagonalisation algorithm (which applies to arbitrary square matrices and can return "not diagonalisable").

Example

Orthogonally diagonalise

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

Step 1. Find the eigenvalues of A: $\lambda_1 = 2$ and $\lambda_2 = 8$.

Find bases for the eigenspaces: $\{\mathbf{u}_1 = (-1, 1, 0), \mathbf{u}_2 = (-1, 0, 1)\}$ for $\lambda_1 = 2$ and $\{\mathbf{u}_3 = (1, 1, 1)\}$ for $\lambda_2 = 8$

Step 2. Apply Gram-Schmidt to $\{\mathbf{u}_1, \mathbf{u}_2\}$ to get an orthonormal basis for $span(\mathbf{u}_1, \mathbf{u}_2)$: $\mathbf{v}_1 = (-1/\sqrt{2}, 1/\sqrt{2}, 0)$ and $\mathbf{v}_2 = (-1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6})$.

Apply G-S to (i.e. normalise) $\mathbf{u}_3 = (1, 1, 1)$ to get $\mathbf{v}_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$.

Step 3. Form the matrix Q. We have

$$Q = \begin{pmatrix} -1\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \quad \text{and} \quad Q^T A Q = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Spectral decomposition

Let A be a symmetric $n \times n$ matrix and let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A, with $\lambda_1, \ldots, \lambda_n$ the corresponding eigenvalues.

$$A = QDQ^T = (\mathbf{u}_1 | \ldots | \mathbf{u}_n) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} =$$

$$(\lambda_1 \mathbf{u}_1 | \ldots | \lambda_n \mathbf{u}_n) \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$$

Theorem (Spectral decomposition)

Let A be a symmetric $n \times n$ matrix. With the $\mathbf{u}_i$'s and the λ_i's as above,

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$$

Geometric interpretation of spectral decomposition

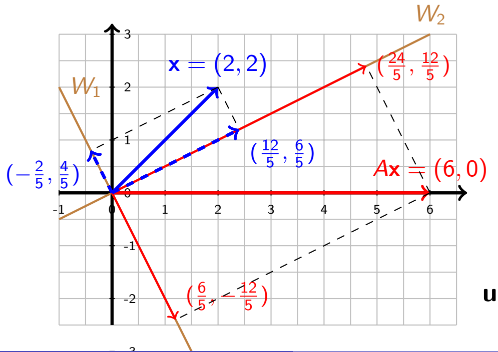
Consider the matrix transformation T_A corresponding to a symmetric matrix A:

$$T_A(\mathbf{x}) = A\mathbf{x} = (\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T)\mathbf{x}.$$

If we denote $W_i = span(\mathbf{u}_i)$, then we have

$$\lambda_i \mathbf{u}_i \mathbf{u}_i^T \mathbf{x} = \lambda_i \mathbf{u}_i (\mathbf{u}_i^T \mathbf{x}) = \lambda_i \mathbf{u}_i (\langle \mathbf{u}_i, \mathbf{x} \rangle) = \lambda_i (\langle \mathbf{u}_i, \mathbf{x} \rangle \mathbf{u}_i) = \lambda_i \operatorname{proj}_{W_i} \mathbf{x}, \text{ and so}$$

$$T_A(\mathbf{x}) = A\mathbf{x} = \lambda_1 \operatorname{proj}_{W_1} \mathbf{x} + \lambda_2 \operatorname{proj}_{W_2} \mathbf{x} + \ldots + \lambda_n \operatorname{proj}_{W_n} \mathbf{x}.$$



$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$$

$$\lambda_1 = -3, \quad \lambda_2 = 2$$

$$\mathbf{u}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \mathbf{u}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

What we learnt today

- Orthogonal matrices, their characterisations and properties
- Spectral theorem: Orthogonally diagonalisable = symmetric
- Eigen-properties of symmetric matrices
- Spectral decomposition for symmetric matrices

Next time:

- Singular value decomposition