Takagi Factorisation

Carl Martin Fevang

September 30, 2023

Abstract

This document describes the diagonalisation procedure by Takagi for diagonalising symmetric, complex-valued matrices.

1 Schur decomposition and singular-value decomposition

Schur decomposition tells us that any (potentially complex) matrix A can be written as

$$A = U^\dagger \Delta U,$$

where U is a unitary matrix, and Δ is an upper triangular matrix. It follows then that if A as a symmetric matrix ($A^T = A$), then

$$\left(U^\dagger \Delta U\right) = \left(U^\dagger \Delta U\right)^T = U^T \Delta^T U^*$$

2 Takagi factorisation

Assume $A = A^T$ is a symmetric, complex-valued, $n \times n$ matrix. Takagi factorisation[1] tells us that there exists a unitary matrix U, and a real, non-negative diagonal matrix D such that

$$A = U^T D U. \tag{1}$$

2.1 Factorisation algorithm

The algorithm is will be based on finding vector $\boldsymbol{v} \in \mathbb{C}^n$ that satisfy $A\boldsymbol{v}^* = \sigma \boldsymbol{v}$, for some real, positive σ. This vector will be called a *Takagi vector* for future reference. Existence of these vectors for any matrix A such that AA^* only has real, positive eigenvalues is detailed later.

To find U, I propose here an algorithm based on the proof for Takagi factorisation in.[2] Given a Takagi vector $\boldsymbol{v} \in \mathbb{C}^n$ of A,[3] and an orthonormal basis $\{\boldsymbol{v}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ of $\mathbb{C}^n$, it is possible to write A as

$$A = V \begin{bmatrix} \sigma & \mathbf{0} \\ \mathbf{0} & A_2 \end{bmatrix} V^T,$$

[1] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990. ISBN: 0521386322. URL: `http://www.amazon.com/Matrix-Analysis-Roger-Horn/dp/0521386322%3FSubscriptionId%3D192BW6DQ43CK9FN0ZGG2%26tag%3Dws%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0521386322`.

[2] Horn and Johnson, *Matrix Analysis*.

[3] A proof that this can be found is detailed elsewhere.

where A_2 is a symmetric $(n-1)\times(n-1)$ matrix and V is a unitary matrix with the aforementioned orthonormal basis as its columns. This process can be repeated with A_2 and so on until you have

$$A = V_1 \cdots V_n \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix} V_n^T \cdots V_1^T,$$

where

$$V_p = \begin{bmatrix} \mathbb{I}_{(p-1)\times(p-1)} & \mathbf{0} \\ \mathbf{0} & \tilde{V}_p, \end{bmatrix}$$

and $\tilde{V}_p$ is the unitary matrix that makes a diagonalisation step on A_p. Comparing to Eq. (1), we find that

$$U = V_n^T \cdots V_1^T, \tag{2a}$$
$$D = \operatorname{diag}(\sigma_1, \ldots, \sigma_n). \tag{2b}$$

It is easy to show that U is unitary, as promised. Furthermore, by assumption, all the values σ_p are real and positive. Now the values on the diagonal of D can be permuted to any order using a permutation matrix P, such that we get

$$A = U_P^T D_P U_P,$$

where $U_P = PU$ and $D_P = PDP^T$. U_P is still unitary, and D_P diagonal.

3 Proofs

The Takagi vector. For any $A \in M_n(\mathbb{C})$ such that AA^* only has real, positive eigenvalues, there exists a vector $\boldsymbol{v} \in \mathbb{C}^n$ such that $A\boldsymbol{v}^* = \sigma\boldsymbol{v}$, where σ is a real, positive number.
Proof. Consider a vector $\boldsymbol{x} \in \mathbb{C}^n$ that is an eigenvector of AA^* with corresponding eigenvalue λ. There are two cases:

(a) $A\boldsymbol{x}^*$ and $\boldsymbol{x}$ are linearly dependent.

(b) $A\boldsymbol{x}^*$ and $\boldsymbol{x}$ are linearly independent.

In case (a), we must have that $A\boldsymbol{x}^* = \mu\boldsymbol{x}$ for some $\mu \in \mathbb{C}$, since they are linearly dependent. Then $AA^*\boldsymbol{x} = A\mu^*\boldsymbol{x}^* = |\mu|^2\,\boldsymbol{x} \equiv \lambda\boldsymbol{x}$, which is positive by definition.
In case (b), the vector $\boldsymbol{y} = A\boldsymbol{x}^* + \mu\boldsymbol{x}$ is non-zero for any $\mu \in \mathbb{C}$, since $A\boldsymbol{x}^*$ and $\boldsymbol{x}$ are linearly independent. Then we can choose μ such that $|\mu|^2 = \lambda$ to get that $A\boldsymbol{y}^* = A\left(A^*\boldsymbol{x} + \mu^*\boldsymbol{x}^*\right) = \lambda\boldsymbol{x} + \mu^*A\boldsymbol{x}^* = \mu\mu^*\boldsymbol{x} + \mu^*A\boldsymbol{x}^* = \mu^*\left(A\boldsymbol{x}^* + \mu\boldsymbol{x}\right) = \mu^*\boldsymbol{y}$.
As such, we can always find a vector $\tilde{\boldsymbol{v}} \in \mathbb{C}^n$ such that $A\tilde{\boldsymbol{v}}^* = \mu\tilde{\boldsymbol{v}}$ for some $\mu \in \mathbb{C}^n$. Furthermore, we can define a vector $\boldsymbol{v} = e^{i\theta}\tilde{\boldsymbol{v}}$ for a $\theta \in \mathbb{R}$ to get $A\boldsymbol{v}^* = A\left(e^{i\theta}\tilde{\boldsymbol{v}}\right)^* = e^{-i\theta}A\tilde{\boldsymbol{v}}^* = e^{-i\theta}\mu\tilde{\boldsymbol{v}} = e^{-2i\theta}\mu e^{i\theta}\tilde{\boldsymbol{v}} = e^{-2i\theta}\mu\boldsymbol{v} \equiv \sigma\boldsymbol{v}$. This allows us to choose the phase of $\sigma = e^{-2i\theta}\mu$ to be such that σ is real and positive.

Eigenvalues of AA^* for symmetric A. Given an $N\times N$ complex matrix A, the eigenvalues of AA^* are always real and non-negative.
Proof. Consider the eigenvectors $\{\boldsymbol{x}_i\}$ of A with corresponding eigenvalues μ_i. Since A is symmetric, these eigenvectors form a basis for $\mathbb{C}^n$. As such, any vector $\boldsymbol{v} \in \mathbb{C}^n$ can be written as $\boldsymbol{v} = \sum_i a_i\boldsymbol{x}_i$ for some coefficients $a_i \in \mathbb{C}$. The eigenvalues of AA^*

are real and non-negative if and only if AA^* is positive semi-definite. To check this, we can look at

$$
\begin{aligned}
\boldsymbol{v}^T AA^* \boldsymbol{v}^* &= \left(\sum_i A^T a_i \boldsymbol{x}_i\right)^T \left(\sum_j A^* a_j^* \boldsymbol{x}_j^*\right) = \sum_{ij} a_i a_j^* \left(A\boldsymbol{x}_i\right)^T \left(A^* \boldsymbol{x}_j^*\right) \\
&= \sum_{ij} a_i a_j^* \left(\mu_i \boldsymbol{x}_i\right)^T \left(\mu_j^* \boldsymbol{x}_j^*\right) = \sum_{ij} a_i a_j^* \mu_i \mu_j^* \left(\boldsymbol{x}_i^\dagger \boldsymbol{x}_j\right)^* \\
&= \sum_i \left\{ |a_i|^2 |\mu_i|^2 \|x_i\|^2 + 2 \sum_{j>i} \operatorname{Re}\left\{ a_i a_j^* \mu_i \mu_j^* \left(\boldsymbol{x}_i^\dagger \boldsymbol{x}_j\right)^* \right\} \right\} \geqslant 0,
\end{aligned}
$$

where we have used that A is symmetric in the second transition. Since any vector in $\mathbb{C}^n$ can be written on the form of $\boldsymbol{v}$, this means that AA^* must be positive semi-definite, and all its eigenvalues must be real and non-negative. This last part can easily be seen by considering an eigenvector $\boldsymbol{y}$ of AA^* with corresponding eigenvalue λ. Then

$$
0 \leqslant \boldsymbol{y}^\dagger AA^* \boldsymbol{y} = \boldsymbol{y}^\dagger \lambda \boldsymbol{y} = \lambda \|\boldsymbol{y}\|^2,
$$

meaning $\lambda \geqslant 0$.

$$
AA^* = AA^\dagger = (AA^\dagger)^\dagger
$$

$$
\boldsymbol{x}^T AA^* \boldsymbol{x}^* = (A^T \boldsymbol{x})^T (A\boldsymbol{x})^* = (A\boldsymbol{x})^T (A\boldsymbol{x})^* = \mu \boldsymbol{x}^T \mu^* \boldsymbol{x}^* = |\mu|^2 (\boldsymbol{x}^\dagger \boldsymbol{x})^* \geqslant 0
$$

$$
\sum_i a_i^* \boldsymbol{x}_i^T AA^* \sum_j a_j \boldsymbol{x}_j^* = \sum_{ij} a_i^* a_j (A\boldsymbol{x}_i)^T (A\boldsymbol{x}_j)^* = \sum_{ij} a_i^* a_j \mu_i \mu_j^* (\boldsymbol{x}_i^\dagger \boldsymbol{x}_j)^*
$$

$$
= \sum_i |a_i|^2 |\mu_i|^2 \|\boldsymbol{x}_i\|^2 + 2 \sum_{j>i} \operatorname{Re}\left\{ a_i^* a_j \mu_i \mu_j^* (\boldsymbol{x}_i^\dagger \boldsymbol{x}_j)^* \right\}
$$

$$
(V^\dagger A V^*)_{i0} = V_{ki}^* A_{kl} V_{l0}^* = V_{ki}^* \sigma V_{k0} = \sigma \delta_{i0}
$$

$$
(V^\dagger A V^*)_{0i} = V_{k0}^* A_{kl} V_{li}^* = A_{lk} V_{k0}^* V_{li}^* = \sigma V_{l0} V_{li}^* = \sigma \delta_{0i}
$$

$$
\boldsymbol{v}_i^\dagger A \boldsymbol{v}^* = \boldsymbol{v}_i^\dagger \sigma \boldsymbol{v} = \sigma \boldsymbol{v}_i^\dagger \boldsymbol{v} = \sigma \boldsymbol{e}_0
$$

$$
\boldsymbol{v}^\dagger A \boldsymbol{v}_i = (A^T \boldsymbol{v}^*)^T \boldsymbol{v}_i = (A\boldsymbol{v}^*)^T \boldsymbol{v}_i = \sigma \boldsymbol{v}^\dagger \boldsymbol{v}_i = \sigma \boldsymbol{e}_0^T
$$

$$
(V^\dagger A V^*)^T = V^\dagger A^T V^* = V^\dagger A V^*
$$

References

Horn, Roger A. and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990. ISBN: 0521386322. URL: `http://www.amazon.com/Matrix-Analysis-Roger-Horn/dp/0521386322%3FSubscriptionId%3D192BW6DQ43CK9FN0ZGG2%26tag%3Dws%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0521386322`.