(ENKEMNA0302) Applied Linear Algebra

Dr. Gabor FACSKO, PhD
Senior Research Fellow
facskog@gamma.ttk.pte.hu

University of Pecs, Faculty of Sciences, Institute of Mathematics and Informatics, 7624 Pecs, Ifjusag utja 6.
Wigner Research Centre for Physics, Department of Space Physics and Space Technology, 1121 Budapest, Konkoly-Thege Miklos ut 29-33.
https://facsko.ttk.pte.hu

March 26, 2025

Trace of a Matrix I

- Definition: (Trace of a matrix). The sum of the elements on the main diagonal of a square matrix is called the trace of the matrix. The trace of matrix $\mathbf{A}$ is denoted by *trace* $\mathbf{A}$ or *tr* $\mathbf{A}$. For example:

$$trace \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 5, \; trace\,(\mathbf{E}_n) = n, \; [\mathbf{a}]_X = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} = 0.$$

- Thesis: (Trace as a linear mapping). The trace is additive and homogeneous, meaning that for any matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ and any scalar $\lambda \in \mathbb{R}$:

$$trace\,(\mathbf{A} + \mathbf{B}) = trace\,(\mathbf{A}) + trace\,(\mathbf{B})\,, trace\,(\lambda \mathbf{A}) = \lambda trace\,(\mathbf{A})\,.$$

Deduction: Trivial. Furthermore, for any constants $\lambda, \mu \in \mathbb{R}$:

$$trace\,(\lambda \mathbf{A} + \mu \mathbf{B}) = \lambda trace\,(\mathbf{A}) + \mu trace\,(\mathbf{B})\,.$$

Trace of a Matrix II

- Furthermore, $trace\left(\mathbf{A}^T\right) = trace\left(\mathbf{A}\right)$.
- Thesis: (Properties of the trace). Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ and $\mathbf{C} \in \mathbb{R}^{m\times n}$. Then:

$$\begin{aligned} trace\left(\mathbf{AB}\right) &= trace\left(\mathbf{BA}\right), \\ trace\left(\mathbf{C}^T\mathbf{C}\right) &= \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^2. \end{aligned}$$

Deduction: Since

$$trace\left(\mathbf{AB}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{n} b_{ij} a_{ji} = trace\left(\mathbf{BA}\right).$$

The proof of the second equality is trivial.

Trace of a Matrix III

- For any two square matrices, $\mathit{trace}\,(\mathbf{AB} - \mathbf{BA}) = 0$.
- The squared length of the vector $\mathbf{x}$ is given by $\mathbf{x} \cdot \mathbf{x} = \mathbf{x}^T\mathbf{x} = \sum_i x_i^2$. This generalizes the second statement of the theorem above.

Diagonalization I

- Definition: (Similarity). We say that the $n \times n$ matrix $\mathbf{A}$ is similar to the matrix $\mathbf{B}$ if there exists an invertible matrix $\mathbf{C}$ such that $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$. Notation: $\mathbf{A} \sim \mathbf{B}$.
- Thesis: (Similarity-invariant properties). If $\mathbf{A}$ and $\mathbf{B}$ are similar matrices, i.e., $\mathbf{A} \sim \mathbf{B}$, then:
  1. $\rho(\mathbf{A}) = \rho(\mathbf{B})$,
  2. $dim(\mathbb{N}(\mathbf{A})) = dim(\mathbb{N}(\mathbf{B}))$,
  3. $\det(\mathbf{A}) = \det(\mathbf{B})$,
  4. $trace(\mathbf{A}) = trace(\mathbf{B})$.

  Deduction:
  1. $\rho(\mathbf{A}) = \rho\left(\mathbf{C}^{-1}\mathbf{B}\mathbf{C}\right) \leq \rho(\mathbf{B})$ and $\rho(\mathbf{B}) = \rho\left(\mathbf{C}^{-1}\mathbf{A}\mathbf{C}\right) \leq \rho(\mathbf{A})$. Thus, $\rho(\mathbf{A}) = \rho(\mathbf{B})$.
  2. $dim(\mathbb{N}(\mathbf{A})) = n - \rho(\mathbf{A}) = n - \rho(\mathbf{B}) = dim(\mathbb{N}(\mathbf{B}))$.
  3. $\det(\mathbf{A}) = \det\left(\mathbf{C}^{-1}\mathbf{B}\mathbf{C}\right) = \det\left(\mathbf{C}^{-1}\right)\det(\mathbf{B})\det(\mathbf{C}) = \det(\mathbf{B})$, since $\det(\mathbf{C})\det\left(\mathbf{C}^{-1}\right) = 1$.

Diagonalization II

4. $trace\left(\mathbf{A}\right) = trace\left(\mathbf{C}^{-1}\mathbf{BC}\right) = trace\left(\mathbf{BCC}^{-1}\right) = trace\left(\mathbf{B}\right)$, using the fact that the trace of a product of matrices remains unchanged under cyclic permutations.

- Definition: (Quadratic form). A real quadratic form (or quadratic function) is a function $\mathbb{R}^n \to \mathbb{R}; \mathbf{x} \to \mathbf{x}^T\mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a real symmetric matrix. A complex quadratic form is defined as the function $\mathbb{C}^n \to \mathbb{C}; \mathbf{x} \to \mathbf{x}^T\mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is a complex square matrix.
- Principal axis theorem, principal axis transformation, quadratic forms and matrix definiteness, determining definiteness from eigenvalues, positive (semi-)definite matrix factorizations, Cholesky decomposition, definiteness and principal minors, extrema problems... Optional: Wettl notes.

Diagonalization III

- Thesis: (Eigenvalue-related invariants). If $\mathbf{A} \sim \mathbf{B}$, then $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic polynomial, and thus the same eigenvalues, including their algebraic and even geometric multiplicities.
  Deduction: Assume that $\mathbf{A} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C}$ for some invertible matrix $\mathbf{C}$. Then:

$$\mathbf{A} - \lambda\mathbf{E} = \mathbf{C}^{-1}\mathbf{B}\mathbf{C} - \lambda\mathbf{C}^{-1}\mathbf{E}\mathbf{C} = \mathbf{C}^{-1}\left(\mathbf{B} - \lambda\mathbf{E}\right)\mathbf{C},$$

  which means that $\mathbf{A} - \lambda\mathbf{E}$ and $\mathbf{B} - \lambda\mathbf{E}$ are similar. Since similar matrices have the same determinant, we obtain $\det\left(\mathbf{A} - \lambda\mathbf{E}\right) = \det\left(\mathbf{B} - \lambda\mathbf{E}\right)$, implying that $\mathbf{A}$ and $\mathbf{B}$ share the same characteristic polynomial, eigenvalues, and algebraic multiplicities. The equality of geometric multiplicities follows from the fact that the nullities of $\mathbf{A} - \lambda\mathbf{E}$ and $\mathbf{B} - \lambda\mathbf{E}$ are equal, as previously proved.

Diagonalization IV

- Definition: (Diagonalizability). The $n \times n$ matrix $\mathbf{A}$ is diagonalizable if it is similar to a diagonal matrix, i.e., if there exists a diagonal matrix $\mathbf{\Lambda}$ and an invertible matrix $\mathbf{C}$ such that $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$.
- Thesis: (Necessary and sufficient condition for diagonalizability). The $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if it has n linearly independent eigenvectors. In this case, the diagonal matrix consists of the eigenvalues of $\mathbf{A}$, and $\mathbf{C}$ is formed from the corresponding eigenvectors.

  Deduction: If $\mathbf{A}$ is similar to a diagonal matrix, there exists a matrix $\mathbf{C}$ such that $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$. Multiplying by $\mathbf{C}$ from the left, we get $\mathbf{C}\mathbf{\Lambda} = \mathbf{A}\mathbf{C}$. Writing $\mathbf{C}$ as $[\mathbf{x}_1\mathbf{x}_2 \ldots \mathbf{x}_n]$ and $\mathbf{\Lambda} = diag\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$, we obtain

Diagonalization V

$$[\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n]\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \lambda_n\end{pmatrix} = \mathbf{A}\,[\mathbf{x}_1\mathbf{x}_2\ldots\mathbf{x}_n]\,.$$

The i-th column of the matrix on the left-hand side is $\lambda_i\mathbf{x}_i$, while the i-th column of the matrix on the right-hand side is $\mathbf{A}\mathbf{x}_i$. These are equal, i.e., $\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i$, so $\mathbf{x}_i$ is an eigenvector corresponding to the eigenvalue λ_i. Since $\mathbf{C}$ is invertible, its column vectors are linearly independent, which proves one part of our statement. Now, suppose that $\mathbf{A}$ has n linearly independent eigenvectors. We construct a diagonal matrix $\mathbf{\Lambda}$ from the eigenvalues, ensuring that the eigenvalue λ_i corresponding to the vector $\mathbf{x}_i$, which is placed in the i-th column of $\mathbf{C}$, is also placed in the i-th column of $\mathbf{\Lambda}$. Since $\lambda_i\mathbf{x}_i = \mathbf{A}\mathbf{x}_i$, it follows that $\mathbf{\Lambda}$ is similar to $\mathbf{A}$.

Diagonalization VI

- The expression $\mathbf{\Lambda} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$ can be rewritten as $\mathbf{A} = \mathbf{C}\mathbf{\Lambda}\mathbf{C}^{-1}$, which is called the eigen decomposition of the matrix $\mathbf{A}$.
- Left eigenvectors and the dyadic form of the eigen decomposition, polynomials of diagonalizable matrices and the Cayley–Hamilton theorem, polynomial of a diagonalizable matrix, Cayley–Hamilton theorem, eigenvectors of distinct eigenvalues, distinct eigenvalues and diagonalizability, the relationship between algebraic and geometric multiplicities, diagonalizability and geometric multiplicity, spectral decomposition of diagonalizable matrices, spectral decomposition of diagonalizable matrices, direct sum of subspaces, properties of the direct sum, eigenspaces of diagonalizable matrices. . . See in Wettl notes.

LU Decomposition

- Definition: (LU Decomposition). We say that an $m \times n$ matrix $\mathbf{A}$ has an LU decomposition (LU factorization or LU decomposition) if it can be written as $\mathbf{A} = \mathbf{LU}$, where $\mathbf{L}$ is a lower unit triangular matrix (i.e., ones on the main diagonal and zeros above it) and $\mathbf{U}$ is an upper triangular matrix.
- Not every matrix has an LU decomposition.
- The LU decomposition is not unique.
- However, it can be shown that if $\mathbf{A}$ is invertible and has an LU decomposition, then it is unique.
- Example of computing an LU decomposition, Algorithm for constructing an LU decomposition, Existence and uniqueness of the LU decomposition, Solving a system of equations using LU decomposition, + example, Inverting a matrix using LU decomposition, + example... in practice.

The End

Thank you for your attention!