Cornell University

Preliminaries: Notation and Basic Algebra

ECE5260/ECE7260/ORIE5735

Anna Scaglione

Cornell Tech

January 21, 2026

Content

Cornell University

Outline

Cornell University

Why these notes?

The goal of these notes is:

- to help you navigate the notation used in the slides
- remind you a few concepts of linear algebra that are constantly used in the course

Outline

Cornell University

Sets

enumerate & itemize

Cornell University

1. **Sets**: I use calligraphic letters for sets and two refer to a graph, which is always $\mathcal{G}$, with $\mathcal{V}$ the set of vertexes/nodes of the graph and $\mathcal{E}$ the set of edges that connect them.
2. **Number of elements of a set**: Is called cardinality, and I refer to it as $|\mathcal{A}|$, though I may use a letter, e.g. $N = |\mathcal{V}|$, and clarify what it is before using it.
3. Sets can contain anything, so to refer to for instance a set of vertices I may use the notation $\mathcal{V} = \{v_1, \ldots, v_N\}$. Sometimes I use the indexes instead $\mathcal{V} = \{1, \ldots, N\}$

Indexes

1. **Indexes of the nodes**: Most commonly the letter v or u are used to refer to nodes.
2. **Indexes of the edges**: For the edges (u, v) is from $v \to u$. Sometimes is conventient to have an interger index $e \in \{1, \ldots, |\mathcal{E}|\}$ to refer to the edge.

Vectors and Matrices

Cornell University

- **Vectors** They are boldface lowercase letters, like $\boldsymbol{x}$.
- If they are a signal with multiple variables (multivariate), I use the notation $\boldsymbol{x}(t)$.
- The entries are either x_v or sometimes is useful to use this notation $[\boldsymbol{x}]_v$ instead of x_v.
- **Matrices**: I use uppecase boldface letters for matrices $\boldsymbol{A}$. I use $a_{u,v}$ for the u, v entry, or also $[\boldsymbol{A}]_{u,v}$
- I tipically consider column vectors, unless otherwise specified:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

Vectors and Matrices

Cornell University

1. An $M \times N$ matrix can be multiplied by an $N \times 1$ vector, with the row by column product and $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ is an $M \times 1$ vector:

$$\boldsymbol{A} = \begin{bmatrix} a_{1,1} & \dots & a_{1,N} \\ \vdots & \vdots & \vdots \\ a_{M,1} & \dots & a_{M,N} \end{bmatrix},$$

$$\boldsymbol{y} = \boldsymbol{A} \;\Rightarrow\; y_u = [\boldsymbol{A}\boldsymbol{x}]_u = \sum_{n=1}^{N} a_{u,n} x_n$$

2. The notation $\boldsymbol{x}^\top$ and $\boldsymbol{A}^\top$ are the transpose of the vector $\boldsymbol{x}$ and matrix $\boldsymbol{A}$ respectively:

$$\boldsymbol{x}^\top = [x_1, \dots, x_N], \quad \boldsymbol{A}^\top = \begin{bmatrix} a_{1,1} & \dots & a_{1,M} \\ \vdots & \vdots & \vdots \\ a_{1,N} & \dots & a_{M,N} \end{bmatrix}$$

Matrix vector products

Cornell University

- **Row by column product**, we already looked at the row by column product of matrices and vectors. We actually have not only $\boldsymbol{A}\boldsymbol{x}$ but also $\boldsymbol{z} = \boldsymbol{w}^\top \boldsymbol{A}$ which is a row vector $(1 \times N)$:

$$z_v = [\boldsymbol{w}^\top \boldsymbol{A}]_v = \sum_{m=1}^{M} a_{m,v} w_m$$

- Let matrix $\boldsymbol{A}$ a number of columns N equal to the number of rows of matrix $\boldsymbol{B}$. The row-by column matrix product is:

$$\boldsymbol{A}\boldsymbol{B}, \quad \Rightarrow \quad [\boldsymbol{A}\boldsymbol{B}]_{uv} = \sum_{n=1}^{N} a_{u,n} b_{n,v}$$

- $(\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top$. Note that the row by column product is not commutative.

Vector products

Cornell University

- **Inner product** $\boldsymbol{a}$ which is $N \times 1$ and $\boldsymbol{b}$ which must also be $N \times 1$, is a scalar number

$$\boldsymbol{a}^\top \boldsymbol{b} = \sum_{n=1}^{N} a_n b_n,$$

and when $\boldsymbol{a}^\top \boldsymbol{b} = 0$ the vectors are orthogonal.

- **Outer product** The outer product of two vectors $\boldsymbol{a}$ which is $N \times 1$ and $\boldsymbol{b}$ which is $M \times 1$ is an $N \times M$ matrix, with entries:

$$\boldsymbol{a}\boldsymbol{b}^\top : \;\Rightarrow\; [\boldsymbol{a}\boldsymbol{b}^\top]_{uv} = a_u b_v$$

Norms

Cornell University

- **Frobenious/Euclidean norm** let $\boldsymbol{x}$ be an $N \times 1$ vector:

$$\boldsymbol{x}^\top \boldsymbol{x} = \|\boldsymbol{x}\|^2 = \sum_{n=1}^{N} x_n^2 \geq 0$$

it is the square of the norm $\|\boldsymbol{x}\|$ of the vector.

- In general we call ℓ_p norm the following: $\|\boldsymbol{a}\|_p = \sqrt[p]{\sum_{n=1}^N |a_n|^p}$
- The Euclidean norm is ℓ_2 norm, $\|\boldsymbol{a}\|_2 = \sqrt[2]{\sum_{n=1}^N |a_n|^2}$
- Interesting are the ℓ_0 and ℓ_1 norms:

$$\|\boldsymbol{a}\|_0 = \text{number of entries } a_n \neq 0$$

which is also called the *support* of vector $\boldsymbol{a}$. And

$$\|\boldsymbol{a}\|_1 = \sum_{n=1}^{N} |a_n|$$

Special matrix and vectors

- $\boldsymbol{I}$ is the identity matrix, $\mathbf{0}$ is the all zero vector and $\mathbf{1}$ the all one vector:

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 1 & 0 \\ 0 & \ddots & \ddots & 0 & 1 \end{bmatrix} \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Outline

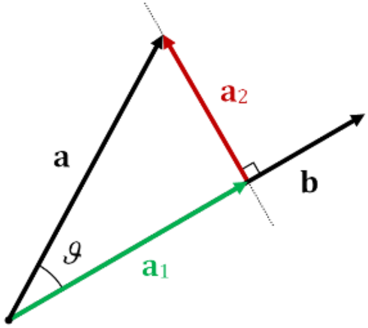
Cornell University

Inner product and projection

- Let $\|\boldsymbol{a}\|$ and $\|\boldsymbol{b}\|$ are the norms of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ respectively (see next slide).
- Let ϑ be the angle between $\|\boldsymbol{a}\|$ and $\|\boldsymbol{b}\|$. The innerproduct is also tied to the so called orthogonal projection of $\boldsymbol{a}$ onto $\boldsymbol{b}$, $\boldsymbol{a}_1$:

$$\boldsymbol{a}_1 = \|\boldsymbol{a}\| \cos(\vartheta) \cdot \frac{\boldsymbol{b}}{\|\boldsymbol{b}\|}$$

- Note that $\boldsymbol{a}_2 = \boldsymbol{a} - \boldsymbol{a}_1$ is orthogonal to $\boldsymbol{b}$ and $\|\boldsymbol{a}_2| = \|\boldsymbol{a}\| \sin(\vartheta)$



$$\boldsymbol{a}^\top \boldsymbol{b} = \|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos(\vartheta), \quad \Rightarrow \quad |\boldsymbol{a}^\top \boldsymbol{b}| = \|\boldsymbol{a}\|\|\boldsymbol{b}\|| \cos(\vartheta)| \equiv \pm\|\boldsymbol{a}_1\|,$$

Determinant and trace

Cornell University

- Associated to square matrices have two important scalar numbers:
  1. **trace** $\mathrm{Tr}(\boldsymbol{A}) = \sum_{n=1}^{N} a_{n,n}$ is the sum of the diagonal elements
  2. **determinant** $|\boldsymbol{A}|$ where we use the Laplacian expansion.
  - A minor $\boldsymbol{M}_{ij}$ of $\boldsymbol{A}$ is the $(N-1) \times (N-1)$ matrix obtain by removing the elements of the ith row and jth column. You can take any row index i and compute the determinant as:

$$|\boldsymbol{A}| = \sum_{n=1}^{N} (-1)^{i+j} a_{ij} |\boldsymbol{M}_{ij}|$$

then you repeat the process for every $|\boldsymbol{M}_{ij}|$ till you get a decomposition in 2×2 minors where we know:

$$\boldsymbol{M} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad |\boldsymbol{M}| = ad - bc.$$

Properties of trace

- $\mathrm{Tr}(a) = a$, where a is a scalar number.
- Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square. For the trace we can write:

$$\mathrm{Tr}(\boldsymbol{AB}) = \mathrm{Tr}(\boldsymbol{BA})$$

$$\Rightarrow \quad \mathrm{Tr}(\boldsymbol{M}^{-1}\boldsymbol{AM}) = \mathrm{Tr}(\boldsymbol{AMM}^{-1}) = \mathrm{Tr}(\boldsymbol{A})$$

$$\Rightarrow \quad \mathrm{Tr}(\boldsymbol{ab}^\top) = \mathrm{Tr}(\boldsymbol{b}^\top\boldsymbol{a}) = \boldsymbol{b}^\top\boldsymbol{a}$$

- The trace is a linear operator, i.e.:

$$\mathrm{Tr}\left(\sum_{i=1}^{I}\alpha_i\boldsymbol{A}_i\right) = \sum_{i=1}^{I}\alpha_i\mathrm{Tr}\left(\sum_{i=1}^{I}\boldsymbol{A}_i\right)$$

Diagonal matrices

- We use the notation $\text{diag}(\boldsymbol{a})$ to denote a diagonal matrix whose diagonal elements are the entries of the vector $\boldsymbol{a}$.

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \quad \text{diag}(\boldsymbol{a}) = \begin{bmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & a_{N-1} & 0 \\ 0 & \ddots & \ddots & 0 & a_N \end{bmatrix}$$

Diagonal matrices

- The following holds

$$(\mathrm{diag}(\boldsymbol{a}))^n = \begin{bmatrix} a_1^n & 0 & 0 & \dots & 0 \\ 0 & a_2^n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & a_{N-1}^n & 0 \\ 0 & \ddots & \ddots & 0 & a_N^n \end{bmatrix}$$

that holds also when $n < 0$. In particular for $n = -1$ we obtain the inverse, i.e. $(\mathrm{diag}(\boldsymbol{a}))^{-1}(\mathrm{diag}(\boldsymbol{a})) = \boldsymbol{I}$

Diagonal matrices

- Let $\boldsymbol{A} = \text{diag}(\boldsymbol{a})$ and consider a matrix $\boldsymbol{B}$, which is $N \times M$, the rows of $\boldsymbol{AB} = \text{diag}(\boldsymbol{a})\boldsymbol{B}$ are $a_r \boldsymbol{b}_r$, $r = 1, \ldots, N$

$$\boldsymbol{C} = \boldsymbol{AB} = \text{diag}(\boldsymbol{a})\boldsymbol{B} \quad \Rightarrow \quad c_{u,v} = [\text{diag}(\boldsymbol{a})\boldsymbol{B}]_{u,v} = a_u b_{u,v}$$

  basically, all elements in row r is multiplied by a_r
- Similarly, if $\boldsymbol{D}$ is $K \times N$, $\boldsymbol{D}\text{diag}(\boldsymbol{a})$ has columns equal $\boldsymbol{d}_c a_c$, $c = 1, \ldots, N$.

$$\boldsymbol{C} = \boldsymbol{D}\text{diag}(\boldsymbol{a}), \quad \Rightarrow \quad c_{u,v} = [\boldsymbol{D}\text{diag}(\boldsymbol{a})]_{u,v} = d_{u,v} a_v$$

- Finally, it is not hard to verify that:

$$\text{diag}(\boldsymbol{a})\boldsymbol{b} = \text{diag}(\boldsymbol{b})\boldsymbol{a} = \begin{bmatrix} a_1 b_1 \\ \vdots \\ a_N b_N \end{bmatrix}$$

Linear independence

Cornell University

- A set of K vectors $\boldsymbol{v}_k$, $k = 1, \ldots, K$ are linearly independent if none of them is a linear combination of the others, i.e. one cannot find a set of coefficients $\boldsymbol{\alpha} \neq \mathbf{0}$ such that

$$\alpha_1 \boldsymbol{v}_1 + \ldots + \alpha_K \boldsymbol{v}_K \equiv \mathbf{0}$$

- Linear combinations of independent K independent vectors $\boldsymbol{v}_k$ that are $N \times 1$ with $N \leq K$ are said to *span* and linear subspace (or the entire space if $N = K$)
- For example two linearly independent 3×1 vectors span a subspace that is a plane, i.e is a two-dimensional space.

Rank of a matrix

- You can view a matrix $\boldsymbol{A}$, which is $M \times N$ as a set of vectors, i.e. the rows

$$\boldsymbol{a}_r = [a_{r,1} \ldots, a_{r,N}], \quad r = 1, \ldots, M,$$

or the columns

$$\boldsymbol{a}_c = \begin{bmatrix} a_{1,c} \\ \vdots \\ a_{M,c} \end{bmatrix}, \quad c = 1, \ldots, N.$$

- The rank of a matrix $\boldsymbol{A}$, is denoted by $r(\boldsymbol{A})$: it is the maximum between the number of linearly independent columns and linearly independent rows of a matrix.

Rank of a matrix

Cornell University

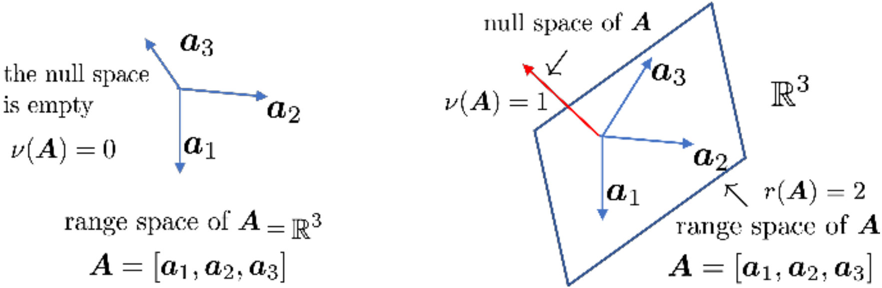
- An $M \times N$ matrix $\boldsymbol{A}$ has a rank

$$r(\boldsymbol{A}) \leq \min(M, N).$$

- It can even one. The rank of the matrix $\boldsymbol{M} = \boldsymbol{a}\boldsymbol{b}^\top$ is actually 1, and any rank one matrix can be written as the outerproduct of two vectors.
- The right null space of a matrix is the subset of vectors $\boldsymbol{v} \neq 0$ such that $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{0}$. The left null-space is $\boldsymbol{v}^\top \boldsymbol{A} = \boldsymbol{0}$
- The dimension of the nullspace is called the nullity $\nu(\boldsymbol{A})$.
- The following is true $r(\boldsymbol{A}) + \nu(\boldsymbol{A}) = N$.

Range space and null space

- We visualize these facts in 3D.



Rank and invertibility of linear equations

Cornell University

- A matrix $\boldsymbol{A}$ can be:
  1. **square**: $\boldsymbol{A}$ is $N \times N$, and its rank is $r(\boldsymbol{A}) \leq N$. If it is $r(\boldsymbol{A}) = N$ the matrix is invertible, and there is a unique solution of $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ i.e.

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}, \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{y}$$

  Note that $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

- Invertible matrices form a **basis** for the Euclidean space $\mathbb{R}^N$, which means that any vector in the space can be expressed as a linear combination of the columns $\boldsymbol{a}_c$ (it is a column vector) or rows $\boldsymbol{a}_r$ (if it is a row vector) of $\boldsymbol{A}$
- A square matrix is full rank if and only if $|\boldsymbol{A}| \neq 0$ and is rank-deficient if and only if $|\boldsymbol{A}| = 0$

Rank of a matrix and invertibility of linear operations

Cornell University

- Also, $\boldsymbol{A}$ can be:
  1. **fat**: $\boldsymbol{A}$ is $M \times N$, where $M < N$.
     - Since $r(\boldsymbol{A}) \leq M < N$ and the M rows of $\boldsymbol{A}$ can span at most an M-dimensional subspace of $\mathbb{R}^N$ the right null space has dimension at least $N - M$.
     - This matrix is not invertible and $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ has multiple (infinite uncountable) solutions. Thus there exist $\boldsymbol{x}' \neq \boldsymbol{x}$ such that $\boldsymbol{A}(\boldsymbol{x}' - \boldsymbol{x}) = \boldsymbol{0}$. In turn $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}'$, so there are multiple solutions, all that differ by a vector in the right null-space of $\boldsymbol{A}$.
  2. **tall**: $\boldsymbol{A}$ is $M \times N$, where $M > N$.
     - This matrix is not invertible and, in this case, technically $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ has no solution without further constraints on $\boldsymbol{x}$.

Orthogonal matrices

Cornell University

- An orthogonal square matrix is such that $\boldsymbol{A}^\top \boldsymbol{A} = \text{diag}([\|\boldsymbol{a}_1\|^2], \ldots, \|\boldsymbol{a}_N\|^2)$, and similarly for $\boldsymbol{A}\boldsymbol{A}^\top$.
- In words the columns are rows are orthogonal (i.e. $\boldsymbol{a}_c^\top \boldsymbol{a}_{c'} = 0$ for all $c \neq c'$) and the diagonal elements of $\boldsymbol{A}^\top \boldsymbol{A}$ are the vectors norms.

Orthonormal matrices

Cornell University

- A matrix $\boldsymbol{A}$ is **orthonormal** if $\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{I}$, $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I}$. In words its columns and rows are orthogonal and their norm is 1.
- For any vector the coefficients of the linear expansion on the orthonomal basis formed by $\boldsymbol{A}$ are the inner products of the vector with the orthogonal basis

$$\boldsymbol{v} = \sum_{c=1}^{N} \overbrace{(\boldsymbol{v}^\top \boldsymbol{a}_c)}^{\text{coefficient}} \boldsymbol{a}_c$$

- A special case of orthonormal matrix is $\boldsymbol{I}$ whose coloums/rows are called the **canonical** basis.
- The multiplication of a vector by an orthonormal basis does not change its norm: $\|\boldsymbol{v}\| = \|\boldsymbol{A}\boldsymbol{v}\|$

Orthogonal Projection

Suppose $\boldsymbol{A}$, $M \times N$ is a tall matrix (i.e. $N < M$). We can define two matrices called *projectors*

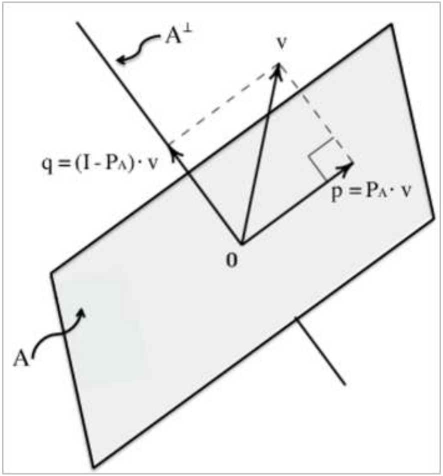
1. The **orthogonal projector**, which can be shown to be

$$\boldsymbol{P_A} = \boldsymbol{A}(\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top,$$

that projects any vector onto the subspace of $\mathbb{R}^M$ spanned by the columns of $\boldsymbol{A}$

2. The projector in the null-space of $\boldsymbol{A}$, which is

$$\boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{I} - \boldsymbol{P_A}.$$



Orthogonal Projection

- The orthogonal projection is the solution of the following linear least square problem:

$$\min_{\boldsymbol{x}} \|\boldsymbol{v} - \boldsymbol{A}\boldsymbol{x}\|^2 \quad \Rightarrow \quad \boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \|\boldsymbol{v} - \boldsymbol{A}\boldsymbol{x}\|^2 = \boldsymbol{P}_A \boldsymbol{v}$$

- The residual error is $\boldsymbol{e} = \boldsymbol{v} - \boldsymbol{P}_A\boldsymbol{v} = (\boldsymbol{I} - \boldsymbol{P}_A)\boldsymbol{v}$
- Hence at the minimum error $\|\boldsymbol{e}\|^2 = \|(\boldsymbol{I} - \boldsymbol{P}_A)\boldsymbol{v}\|^2$ is the norm of the projection of $\boldsymbol{v}$ onto the null-space of $\boldsymbol{A}$.
- Note that both $\boldsymbol{P}_A$ and $\boldsymbol{P}_A^\perp$ are **idenmpotent** matrices, i.e. $\boldsymbol{P}_A\boldsymbol{P}_A = \boldsymbol{P}_A$ and $\boldsymbol{P}_A^\perp\boldsymbol{P}_A^\perp = \boldsymbol{P}_A^\perp$.
- Obviously, $\boldsymbol{P}_A\boldsymbol{P}_A^\perp = \boldsymbol{0}$

Eigenvectors and eigenvalues

- Consider a square matrix $\boldsymbol{A}$ $(N \times N)$ (with some exceptions) we can find N linearly independent vectors $\boldsymbol{v}_n$, $n = 1, \dots, N$ such that:

$$\lambda_n \boldsymbol{v}_n = \boldsymbol{A}\boldsymbol{v}_n, \quad n = 1, \dots, N$$

- The eigenvalues are the solution of the so called characteristic polinomial, which is of Nth order:

$$|\lambda \boldsymbol{I} - \boldsymbol{A}|$$

Eigenvalue Decomposition of $\boldsymbol{A}$

Cornell University

- Let us define the matrix $\boldsymbol{V} = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_N]$ whose columns are the eigenvectors and the vector $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_N]^\top$ whose elements are the eigenvalues of $\boldsymbol{A}$:

$$\boldsymbol{V}\text{diag}(\boldsymbol{\lambda}) = \boldsymbol{A}\boldsymbol{V}^{-1} \quad \Rightarrow \quad \boldsymbol{A} = \boldsymbol{V}\text{diag}(\boldsymbol{\lambda})\boldsymbol{V}$$

- Note that: $\boldsymbol{A}^n = \boldsymbol{V}(\text{diag}(\boldsymbol{\lambda}))^n\boldsymbol{V}^{-1}$ and so:

$$\lim_{n\to+\infty} \boldsymbol{A}^n = \begin{cases} +\infty & \text{if there exist } |\lambda_n| > 1 \\ \mathbf{0} & \text{if all } |\lambda_n| < 1 \end{cases}$$

only if if there exist $|\lambda_n| = 1$ and the remaining are < 1 the entries of limit are finite.

Eigenvalue Decomposition of $\boldsymbol{A}$

Cornell University

- Also a matrix with all non-zero eigenvalues is always invertible

$$\boldsymbol{A}^{-1} = \boldsymbol{V}(\mathrm{diag}(\boldsymbol{\lambda}))^{-1}\ \boldsymbol{V}^{-1}$$

- In fact, the rank of the matrix $r(\boldsymbol{A}) = \|\boldsymbol{\lambda}\|_0$, i.e. the number of non zero eigenvalues of $\boldsymbol{A}$.
- The following holds:

$$|\boldsymbol{A}| = \prod_{n=1}^{N} \lambda_n, \qquad \mathrm{Tr}(\boldsymbol{A}) = \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} a_{nn}$$

Symmetric matrices and eigenvalue decomposition

- A matrix is symmetric if $\boldsymbol{A} = \boldsymbol{A}^\top$. Clearly it must be square.
- Symmetric matrices have an eigenvalue decomposition where the eigenvectors are orhtonormal:

$$\boldsymbol{A} = \boldsymbol{V}\text{diag}(\boldsymbol{\lambda})\boldsymbol{V}^\top,$$

  the preferred notation usually is

$$\boldsymbol{A} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^\top,\ \boldsymbol{V}^\top = \boldsymbol{U}$$

$$\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{I}, \quad \boldsymbol{U}\boldsymbol{U}^\top = \boldsymbol{I}$$

- In general, for any unitary matrix $|\boldsymbol{A}| = |\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\top|$ and $\text{Tr}(\boldsymbol{A}) = \text{Tr}(\boldsymbol{U}\boldsymbol{A}\boldsymbol{U}^\top)$

Properties

Cornell University

- Symmetric matrices are either
  1. **positive definite** iff $\lambda_n > 0, \quad n = 1, \ldots, N,$
     - The for all $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{a}^\top \boldsymbol{A} \boldsymbol{a} > 0$
  2. **positive semi-definite** iff $\lambda_n \geq 0, \quad n = 1, \ldots, N.$
     - The for all $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{a}^\top \boldsymbol{A} \boldsymbol{a} \geq 0$
  3. **negative definite** iff $\lambda_n < 0, \quad n = 1, \ldots, N,$
     - The for all $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{a}^\top \boldsymbol{A} \boldsymbol{a} < 0$
  4. **negative semi-definite** iff $\lambda_n \leq 0, \quad n = 1, \ldots, N.$
     - The for all $\boldsymbol{a} \in \mathbb{R}^N$, $\boldsymbol{a}^\top \boldsymbol{A} \boldsymbol{a} \leq 0$
- Positive and negative definite matrices are always invertible, and in the other cases they are not full rank

Rayleigh quotient

Cornell University

- Assume $\boldsymbol{A}$ is symmetric. The following ratio is called the Rayleigh quotient $\frac{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}}$

Theorem

Let the eigenvalues be in decreasing order $\lambda_1 \geq \ldots \geq \lambda_N$ and $\boldsymbol{v}_n$ denote the eigenvectors. The eigenvalues and eigenvectors are the solutions of the optimizations:

$$\lambda_1 = \max_{\boldsymbol{v}} \frac{\boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}} \tag{1}$$

$$\lambda_2 = \max_{\boldsymbol{v}} \frac{\boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}}, \text{ subject to } \boldsymbol{v}_1^\top \boldsymbol{v} = 0 \tag{2}$$

$$\lambda_N = \max_{\boldsymbol{v}} \frac{\boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v}}{\boldsymbol{v}^\top \boldsymbol{v}}, \text{ subject to } \boldsymbol{v}_n^\top \boldsymbol{v} = 0, n = 1, \ldots, N-1 \tag{3}$$

and the solutions $\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N]$ that maximize the objectives are the eigenvectors. Due to the constraints, $\boldsymbol{V}^\top \boldsymbol{V} = \boldsymbol{I}$.