BASICS OF LINEAR AND MATRIX ALGEBRA

Sergiy A. Vorobyov

Department of Signal Processing and Acoustics
Aalto University

LINEAR AND MATRIX ALGEBRA

Vector signal description

Let the signal is represented by its values $x_1, \dots, x_N$. Then, in vector notation:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{bmatrix}$$

Vector transpose:

$$\mathbf{x}^T = [x_1, x_2, \dots, x_N]$$

Sometimes, it is convenient to consider sets of vectors, for example:

$$\mathbf{x}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \dots \\ x(n-N+1) \end{bmatrix}$$

Vector Euclidean norm:

$$||\mathbf{x}|| = \left\{ \sum_{i=1}^{N} |x_i|^2 \right\}^{1/2}$$

Introducing Hermitian transpose

$$\mathbf{x}^H = \left(\mathbf{x}^T\right)^* = [x_1^*, x_2^*, \dots, x_N^*]$$

we rewrite the norm as

$$||\mathbf{x}|| = \sqrt{\mathbf{x}^H \mathbf{x}}$$

The scalar (inner) product of two complex vectors $\mathbf{a} = [a_1, \ldots, a_N]^T$ and $\mathbf{b} = [b_1, \ldots, b_N]^T$:

$$\mathbf{a}^H \mathbf{b} = \sum_{i=1}^{N} a_i^* b_i$$

Cauchy-Schwarz inequality

$$|\mathbf{a}^H \mathbf{b}| \leq ||\mathbf{a}|| \cdot ||\mathbf{b}||$$

Orthogonal vectors:

$$\mathbf{a}^H \mathbf{b} = \mathbf{b}^H \mathbf{a} = 0$$

The set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is said to be *linearly independent* if

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_n \mathbf{x}_n = 0 \qquad (*)$$

implies that $\alpha_i = 0$ for all i. If any set of nonzero α_i can be found so that $(*)$ holds, then the vectors are *linearly dependent*. For example, for nonzero α_1,

$$\mathbf{x}_1 = \beta_2 \mathbf{x}_2 + \cdots + \beta_n \mathbf{x}_n$$

Example of linearly independent vector set:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Adding to this linearly independent vector set a new vector $\mathbf{x}_3$, we obtain that the new set

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

becomes linearly dependent because

$$\mathbf{x}_1 = \mathbf{x}_2 + 2\mathbf{x}_3$$

Given N vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, consider the set of all vectors that may be formed as a linear combination of the vectors $\mathbf{x}_i$,

$$\mathbf{x} = \sum_{i=1}^{N} \alpha_i \mathbf{x}_i$$

This set forms a *vector space* and the vectors $\mathbf{x}_i$ are said to span this space. If the vectors $\mathbf{x}_i$ are linearly independent, they are said to form a *basis* for this space and the number of basis vectors N is referred to as the space *dimension*. The basis for a vector space is not unique!

Matrices

$n \times m$ matrix:

$$\mathbf{A} = \{a_{ik}\} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nm} \end{bmatrix}$$

Symmetric square matrix:

$$\mathbf{A}^T = \mathbf{A}$$

Hermitian square matrix:

$$\mathbf{A}^H = \mathbf{A}$$

Some properties (apply to transpose $(\cdot)^T$ as well):

$$(\mathbf{A}+\mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H\,, \quad (\mathbf{A}^H)^H = \mathbf{A}\,, \quad (\mathbf{A}\mathbf{B})^H = \mathbf{B}^H\mathbf{A}^H$$

Column and row representations of an $n \times m$ matrix:

$$\mathbf{A} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m] = \begin{bmatrix} \mathbf{r}_1^H \\ \mathbf{r}_2^H \\ \vdots \\ \mathbf{r}_n^H \end{bmatrix} \qquad (*)$$

The *rank* of $\mathbf{A}$ is defined as a number of linearly independent columns in $(*)$, or, equivalently, the number of linearly independent row vectors in $(*)$.

Important property:

$$\mathrm{rank}\{\mathbf{A}\} = \mathrm{rank}\{\mathbf{A}\mathbf{A}^H\} = \mathrm{rank}\{\mathbf{A}^H\mathbf{A}\}$$

For any $n \times m$ matrix:

$$\text{rank}\{\mathbf{A}\} \leq \min\{m, n\}$$

The matrix $\mathbf{A}$ is said to be of *full rank* if

$$\text{rank}\{\mathbf{A}\} = \min\{m, n\}$$

If the square matrix $\mathbf{A}$ is of full rank, then there exists a unique matrix $\mathbf{A}^{-1}$, called the *inverse* of $\mathbf{A}$:

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

The matrix $\mathbf{I}$ is the so-called *identity matrix*:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

The $n \times n$ matrix $\mathbf{A}$ is called *singular* if its inverse does not exist (i.e., if $\mathrm{rank}\{\mathbf{A}\} < n$).

Some properties of inverse:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad (\mathbf{A}^{H})^{-1} = (\mathbf{A}^{-1})^{H}$$

Determinant of a square $n \times n$ matrix (for any i):

$$\det \mathbf{A} = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det \mathbf{A}_{ik}$$

where $\mathbf{A}_{ik}$ is the $(n-1) \times (n-1)$ matrix formed by deleting the ith row and the kth column of $\mathbf{A}$.

Example:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\det \mathbf{A} = a_{11} \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Property: an $n \times n$ matrix $\mathbf{A}$ is *invertible* (nonsingular) if and only if its determinant is nonzero

$$\det \mathbf{A} \neq 0$$

Some additional important properties of determinant:

$$\det\{\mathbf{AB}\} = \det \mathbf{A} \det \mathbf{B}\ , \quad \det\{\alpha \mathbf{A}\} = \alpha^n \det \mathbf{A}$$

$$\det \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}\ , \quad \det \mathbf{A}^T = \det \mathbf{A}$$

Another important function of matrix is *trace*:

$$\mathrm{trace}\{\mathbf{A}\} = \sum_{i=1}^{n} a_{ii}$$

Linear equations

Many practical DSP problems (such as signal modeling, Wiener filtering, etc.) require the solution to a set of linear equations:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n \end{aligned}$$

In matrix notation

$$\mathbf{Ax} = \mathbf{b}$$

Case 1: square matrix $\mathbf{A}$ $(m = n)$. The nature of solution depends upon whether or not $\mathbf{A}$ is singular. In the *nonsingular* case

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

If $\mathbf{A}$ is singular, there may be *no solution* or *many solutions*.

Example:

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2 \qquad \text{no solution} \end{aligned}$$

However, if we modify the equations:

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 1 \qquad \text{many solutions} \end{aligned}$$

Case 2: rectangular matrix $\mathbf{A}$ $(m < n)$. *More equations than unknowns* and, in general, *no solution exist*. The system is called *overdetermined*. In the case when $\mathbf{A}$ is a full rank matrix, and, therefore, $\mathbf{A}^H\mathbf{A}$ is nonsingular, the common approach is to find *least squares solution* by minimizing the norm of the error vector

$$
\begin{aligned}
||\mathbf{e}||^2 &= ||\mathbf{b} - \mathbf{A}\mathbf{x}||^2 \\
&= (\mathbf{b} - \mathbf{A}\mathbf{x})^H(\mathbf{b} - \mathbf{A}\mathbf{x}) \\
&= \mathbf{b}^H\mathbf{b} - \mathbf{x}^H\mathbf{A}^H\mathbf{b} - \mathbf{b}^H\mathbf{A}\mathbf{x} + \mathbf{x}^H\mathbf{A}^H\mathbf{A}\mathbf{x} \\
&= \left[\mathbf{x} - (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{b}\right]^H(\mathbf{A}^H\mathbf{A})\left[\mathbf{x} - (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{b}\right] \\
&+ \left[\mathbf{b}^H\mathbf{b} - \mathbf{b}^H\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{b}\right]
\end{aligned}
$$

The second term is *independent* of $\mathbf{x}$. Therefore, the LS solution is

$$\mathbf{x}_{\mathrm{LS}} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{b}$$

The best (LS) approximation of $\mathbf{b}$ is given by

$$\hat{\mathbf{b}} = \mathbf{A}\mathbf{x}_{\mathrm{LS}} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{b} = \mathbf{P}_{\mathbf{A}}\mathbf{b}$$

where

$$\mathbf{P}_{\mathbf{A}} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$$

is the so-called *projection matrix* with the properties

$$\mathbf{P}_{\mathbf{A}}\mathbf{a} = \mathbf{a}$$

if the vector $\mathbf{a}$ belongs to the column-space of $\mathbf{A}$ and

$$\mathbf{P}_{\mathbf{A}}\mathbf{a} = 0$$

if this vector is orthogonal to the columns of $\mathbf{A}$

The minimum LS error

$$\begin{aligned}
||e||^2_{\min} &= ||\mathbf{b} - \mathbf{A}\mathbf{x}_{\mathrm{LS}}||^2 \\
&= ||(\mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H)\mathbf{b}||^2 \\
&= ||(\mathbf{I} - \mathbf{P}_{\mathbf{A}})\mathbf{b}||^2 = ||\mathbf{P}^{\perp}_{\mathbf{A}}\mathbf{b}||^2 = \mathbf{b}^H\mathbf{P}^{\perp}_{\mathbf{A}}\mathbf{b}
\end{aligned}$$

where $\mathbf{P}^{\perp}_{\mathbf{A}} = \mathbf{I} - \mathbf{P}_{\mathbf{A}}$ is the projection matrix on the subspace orthogonal to the column-space of $\mathbf{A}$.

Alternatively, the LS solution is found from the *normal equations*

$$\mathbf{A}^H\mathbf{A}\mathbf{x} = \mathbf{A}^H\mathbf{b}$$

Case 3: rectangular matrix $\mathbf{A}$ $(n < m)$. *Fewer equations than unknowns* and, provided the equations are consistent, there are *many solutions*. The system is called *underdetermined*.

Special matrix forms

Diagonal square matrix:

$$\mathbf{A} = \mathrm{diag}\{a_{11}, a_{22}, \ldots, a_{nn}\} = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

Exchange matrix:

$$\mathbf{J} = \begin{bmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Toeplitz matrix:

$$a_{ik} = a_{i+1,k+1} \text{ for all } i, k < n$$

Example:

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 2 & 1 & 3 & 2 \\ 7 & 2 & 1 & 3 \\ 1 & 7 & 2 & 1 \end{bmatrix}$$

2.4 Quadratic and Hermitian forms

Quadratic form of a real symmetric square matrix $\mathbf{A}$:

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

Similarly, *Hermitian form* of a Hermitian square matrix $\mathbf{A}$:

$$Q(\mathbf{x}) = \mathbf{x}^H \mathbf{A} \mathbf{x}$$

Symmetric (Hermitian) matrices are positive semidefinite if $Q(\mathbf{x}) \geq 0$ for all nonzero $\mathbf{x}$.

Example: the matrix $\mathbf{A} = \mathbf{y}\mathbf{y}^H$ is positive semidefinite, where $\mathbf{y}$ is an arbitrary complex vector:

$$Q(\mathbf{x}) = \mathbf{x}^H \mathbf{y}\mathbf{y}^H \mathbf{x} = |\mathbf{x}^H \mathbf{y}|^2 \geq 0$$

Eigenvalues and eigenvectors

Consider the *characteristic equation* of an $n \times n$ matrix $\mathbf{A}$:

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

This is equivalent to the following set of *homogeneous linear equations*

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{u} = 0$$

Therefore, the matrix $\mathbf{A} - \lambda\mathbf{I}$ is *singular*. Hence,

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

where $p(\lambda)$ is the so-called *characteristic polynomial* with n roots λ_i $(i = 1, 2 \ldots, n)$ being the *eigenvalues* of $\mathbf{A}$.

For each eigenvalue λ_i, the matrix $\mathbf{A} - \lambda_i \mathbf{I}$ is singular, and, therefore, there will be at least one nonzero *eigenvector* that solves the equation

$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i$$

Since for any eigenvector $\mathbf{u}_i$ any vector $\alpha \mathbf{u}_i$ will be also an eigenvector, the eigenvectors are often *normalized*:

$$||\mathbf{u}_i|| = 1\,, \quad i = 1, 2, \ldots, n$$

Property 1: The eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ corresponding to *distinct* eigenvalues are *linearly independent*.

Property 2: If $\mathrm{rank}\{\mathbf{A}\} = m$, then there will be $n - m$ independent solutions to the homogeneous equation $\mathbf{A}\mathbf{u}_i = 0$. These solutions form the so-called *null-space* of $\mathbf{A}$.

Property 3: The eigenvalues of a Hermitian matrix are *real*.

Proof: From the characteristic equation $\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i$, we have

$$\mathbf{u}_i^H \mathbf{A} \mathbf{u}_i = \lambda_i \mathbf{u}_i^H \mathbf{u}_i \qquad (*)$$

Taking the Hermitian transpose of $(*)$, we have

$$\mathbf{u}_i^H \mathbf{A}^H \mathbf{u}_i = \lambda_i^* \mathbf{u}_i^H \mathbf{u}_i \qquad (**)$$

Since $\mathbf{A}$ is Hermitian $(\mathbf{A} = \mathbf{A}^H)$, $(**)$ becomes

$$\mathbf{u}_i^H \mathbf{A} \mathbf{u}_i = \lambda_i^* \mathbf{u}_i^H \mathbf{u}_i \qquad (***)$$

Finally, comparison of $(*)$ and $(***)$ shows that λ_i are real.

<u>**Property 4:**</u> A Hermitian matrix is *positive definite* if and only if the eigenvalues of $\mathbf{A}$ are *positive*.

Similar property holds for *positive semidefinite*, *negative definite*, or *negative semidefinite* matrices.

A useful *relationship* between matrix determinant and eigenvalues:

$$\det\{\mathbf{A}\} = \prod_{i=1}^{n} \lambda_i$$

Therefore, any matrix is *invertible* (nonsingular) if and only if *all of its eigenvalues are nonzero*.

Property 5: The eigenvectors of a Hermitian matrix corresponding to distinct eigenvalues are *orthogonal*, i.e., if $\lambda_i \neq \lambda_k$, then $\mathbf{u}_i^H \mathbf{u}_k = 0$.

Proof: Let λ_i and λ_k be two *distinct* eigenvalues of $\mathbf{A}$. Then

$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i \quad \text{and} \quad \mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_k$$

Multiplying these equations by $\mathbf{u}_k^H$ and $\mathbf{u}_i^H$, respectively, yields

$$\mathbf{u}_k^H \mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_k^H \mathbf{u}_i \,, \quad \mathbf{u}_i^H \mathbf{A}\mathbf{u}_k = \lambda_k \mathbf{u}_i^H \mathbf{u}_k \qquad (*)$$

Taking the Hermitian transpose of the second equation of $(*)$ and remarking that $\mathbf{A}$ is Hermitian (i.e., $\mathbf{A}^H = \mathbf{A}$ and $\lambda_k^* = \lambda_k$), yields

$$\mathbf{u}_k^H \mathbf{A}\mathbf{u}_i = \lambda_k \mathbf{u}_k^H \mathbf{u}_i \qquad (**)$$

Now, subtracting (∗∗) from the first equation of (∗) leads to

$$0 = (\lambda_i - \lambda_k)\mathbf{u}_k^H \mathbf{u}_i$$

Since the eigenvalues are *distinct* (i.e., $\lambda_i \neq \lambda_k$), we have that

$$\mathbf{u}_k^H \mathbf{u}_i = 0$$

which proofs the *orthogonality* of eigenvectors.

Remark: Although proven above for the distinct eigenvalue case, this property can be *extended* to any $n \times n$ Hermitian matrix with *arbitrary* (not necessarily distinct) eigenvalues.

Eigendecomposition

For an $n \times n$ matrix $\mathbf{A}$, we may perform an *eigendecomposition*:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1} \qquad (*)$$

To do this, let us write the set of equations

$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i\,, \quad i = 1, 2, \ldots, n$$

in the form

$$\mathbf{A}[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n] = [\lambda_1 \mathbf{u}_1, \lambda_2 \mathbf{u}_2, \ldots, \lambda_n \mathbf{u}_n]\,, \quad \text{or, equivalentely}$$

$$\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{\Lambda} \quad \text{with} \quad \mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \qquad (**)$$

and *nonsingular* $\mathbf{U}$. Multiplying $(**)$ on the right by $\mathbf{U}^{-1}$, we get $(*)$.

For a Hermitian matrix, the following property holds because of the orthonormality of eigenvectors:

$$\mathbf{U}^H\mathbf{U} = \mathbf{I}$$

Hence, $\mathbf{U}$ is *unitary* (i.e., $\mathbf{U}^H = \mathbf{U}^{-1}$), and, therefore, the *eigendecomposition* takes the form

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$$

or, equivalently,

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^H$$

Using the unitary property of $\mathbf{U}$, it is easy to find *matrix inverse* via eigendecomposition:

$$\begin{aligned}
\mathbf{A}^{-1} &= (\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H)^{-1} \\
&= (\mathbf{U}^H)^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{U}^{-1} \\
&= \mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^H
\end{aligned}$$

Equivalently

$$\mathbf{A}^{-1} = \sum_{i=1}^{n} \frac{1}{\lambda_i}\mathbf{u}_i\mathbf{u}_i^H$$

Hence, the inverse *does not affect eigenvectors* but *transforms eigenvalues* λ_i to $1/\lambda_i$.

In many applications, matrices may be very close to singular (*ill-conditioned*) and, therefore, their inverse may be *unstable*. We may wish to stabilize the problem by adding a constant to each term along diagonal (the so-called *diagonal loading*):

$$\mathbf{A} = \mathbf{B} + \alpha \mathbf{I}$$

This operation *leaves eigenvectors unchanged* but *changes eigenvalues*:

$$\mathbf{A}\mathbf{u}_i = \mathbf{B}\mathbf{u}_i + \alpha\mathbf{u}_i = (\lambda_i + \alpha)\mathbf{u}_i$$

where λ_i and $\mathbf{u}_i$ are the eigenvalues and eigenvectors of $\mathbf{B}$:

$$\mathbf{B}\mathbf{u}_i = \lambda_i \mathbf{u}_i$$

We can reformulate the trace of $\mathbf{A}$ in terms of eigenvalues:

$$\mathrm{trace}\{\mathbf{A}\} = \sum_{i=1}^{n} \lambda_i \qquad (*)$$

Similarly,

$$\mathrm{trace}\{\mathbf{A}^{-1}\} = \sum_{i=1}^{n} \frac{1}{\lambda_i}$$

This property can be easily proven using the eigendecomposition and the property $\mathrm{trace}\{\mathbf{A}+\mathbf{B}\} = \mathrm{trace}\{\mathbf{A}\} + \mathrm{trace}\{\mathbf{B}\}$. In several applications (such as adaptive filtering), we need some simple and close upper bound for the maximal eigenvalue $\lambda_{\max}$. From $(*)$, we obtain that

$$\lambda_{\max} \leq \mathrm{trace}\{\mathbf{A}\}$$

Singular value decomposition

For a nonsquare $n \times m$ matrix $\mathbf{A}$, we may perform the SVD instead of eigendecomposition:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$$

or, equivalently

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{v}_i^H \quad \text{if } n < m$$

and

$$\mathbf{A} = \sum_{i=1}^{m} \lambda_i \mathbf{u}_i \mathbf{v}_i^H \quad \text{if } n > m$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the $n \times 1$ and $m \times 1$ *left and right singular vectors*, respectively, and λ_i are *singular values*.