# 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  $x_1, \ldots, x_N$ . 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:

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

Vector Euclidean norm:

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

Introducing Hermitian transpose

$$\mathbf{x}^{H} = (\mathbf{x}^{T})^{*} = [x_{1}^{*}, x_{2}^{*}, \dots, x_{N}^{*}]$$

we rewrite the norm as

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

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

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

Cauchy-Schwarz inequality

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

Orthogonal vectors:

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

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

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_n \mathbf{x}_n = 0 \tag{*}$$

implies that  $\alpha_i=0$  for all i. If any set of nonzero  $\alpha_i$  can be found so that (\*) holds, then the vectors are *linearly dependent*. For example, for nonzero  $\alpha_1$ ,

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

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

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

becomes linearly dependent because

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

Given N vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ , 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!

2016

### <u>Matrices</u>

 $n \times m$  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:

$$\mathbf{A} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m] = \left[egin{array}{c} \mathbf{r}_1^H \ \mathbf{r}_2^H \ \vdots \ \mathbf{r}_n^H \end{array}
ight] \qquad (*)$$

The rank of  $\bf A$  is defined as a number of linearly independent columns in (\*), or, equivalently, the number of linearly independent row vectors in (\*). Important property:

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

For any  $n \times m$  matrix:

$$rank{\mathbf{A}} \le min{m, n}$$

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

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

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  ${f I}$  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  $\operatorname{rank}\{\mathbf{A}\} < n$ ).

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:

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

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

Another important function of matrix is trace:

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

## Linear equations

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

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = b_2$$

$$\vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n$$

In matrix notation

$$Ax = b$$

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  $\bf A$  is singular, there may be *no solution* or *many solutions*. Example:

$$x_1 + x_2 = 1$$

$$x_1 + x_2 = 2$$
 no solution

However, if we modify the equations:

$$x_1 + x_2 = 1$$
 $x_1 + x_2 = 1$  many solutions

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

$$||\mathbf{e}||^{2} = ||\mathbf{b} - \mathbf{A}\mathbf{x}||^{2}$$

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

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

$$= \left[\mathbf{x} - (\mathbf{A}^{H}\mathbf{A})^{-1}\mathbf{A}^{H}\mathbf{b}\right]^{H}(\mathbf{A}^{H}\mathbf{A})\left[\mathbf{x} - (\mathbf{A}^{H}\mathbf{A})^{-1}\mathbf{A}^{H}\mathbf{b}\right]$$

$$+ \left[\mathbf{b}^{H}\mathbf{b} - \mathbf{b}^{H}\mathbf{A}(\mathbf{A}^{H}\mathbf{A})^{-1}\mathbf{A}^{H}\mathbf{b}\right]$$

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

$$P_A a = a$$

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

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

if this vector is orthogonal to the columns of  ${f A}$ The minimum LS error

$$||e||_{\min}^{2} = ||\mathbf{b} - \mathbf{A}\mathbf{x}_{LS}||^{2}$$

$$= ||(\mathbf{I} - \mathbf{A}(\mathbf{A}^{H}\mathbf{A})^{-1}\mathbf{A}^{H})\mathbf{b}||^{2}$$

$$= ||(\mathbf{I} - \mathbf{P}_{\mathbf{A}})\mathbf{b}||^{2} = ||\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{b}||^{2} = \mathbf{b}^{H}\mathbf{P}_{\mathbf{A}}^{\perp}\mathbf{b}$$

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

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:

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

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:

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

#### Example:

$$\left[\begin{array}{cccc}
1 & 3 & 2 & 4 \\
2 & 1 & 3 & 2 \\
7 & 2 & 1 & 3 \\
1 & 7 & 2 & 1
\end{array}\right]$$

## 2.4 Quadratic and Hermitian forms

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

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

Similarly, Hermitian form of a Hermitian square matrix  ${f A}$ :

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

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:

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

## 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  ${\bf A}-\lambda {\bf I}$  is singular. Hence,

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

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

For each eigenvalue  $\lambda_i$ , 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*:

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

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

Property 2: If  $rank\{A\} = m$ , 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

$$\mathbf{u}_i^H \mathbf{A} \mathbf{u}_i = \lambda_i \mathbf{u}_i^H \mathbf{u}_i \tag{*}$$

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

$$\mathbf{u}_i^H \mathbf{A}^H \mathbf{u}_i = \lambda_i^* \mathbf{u}_i^H \mathbf{u}_i \tag{**}$$

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

$$\mathbf{u}_i^H \mathbf{A} \mathbf{u}_i = \lambda_i^* \mathbf{u}_i^H \mathbf{u}_i \tag{***}$$

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

Property 4: 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  $\lambda_i$  and  $\lambda_k$  be two *distinct* eigenvalues of **A**. Then

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

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

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

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

$$\mathbf{u}_k^H \mathbf{A} \mathbf{u}_i = \lambda_k \mathbf{u}_k^H \mathbf{u}_i \tag{**}$$

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 **A**, we may perform an eigendecomposition:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{-1} \tag{*}$$

To do this, let us write the set of equations

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

in the form

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

$$\mathbf{AU} = \mathbf{U}\Lambda \quad \text{with} \quad \Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$$
 (\*\*)

and nonsingular U. Multiplying (\*\*) on the right by  $U^{-1}$ , 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 U, it is easy to find *matrix inverse* via eigendecomposition:

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

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  $\lambda_i$  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  $\lambda_i$  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  $\bf A$  in terms of eigenvalues:

$$\operatorname{trace}\{\mathbf{A}\} = \sum_{i=1}^{n} \lambda_i \tag{*}$$

Similarly,

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

This property can be easily proven using the eigendecomposition and the property  $\operatorname{trace}\{\mathbf{A}+\mathbf{B}\}=\operatorname{trace}\{\mathbf{A}\}+\operatorname{trace}\{\mathbf{B}\}.$  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

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

## 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  $\lambda_i$  are singular values.