LINEAR ALGEBRA
Appendix A in Stoica & Moses
Ch. 2.3 in Hayes

Here, we follow Ch. 2.3 in Hayes.

2.3.1 Vectors

Let signal be represented by scalar values $x_1, x_2, \ldots, x_N$. Then, the vector notation is

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}.$$

Vector transpose:

$$\boldsymbol{x}^T = [x_1, x_2, \ldots, x_N].$$

Hermitian transpose:

$$\boldsymbol{x}^H = (x^T)^* = [x_1^*, x_2^*, \ldots, x_N^*].$$

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

$$\boldsymbol{x}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-N+1) \end{bmatrix}.$$

Note: Stoica and Moses use "*" to denote the Hermitian transpose.

Magnitude of a vector?

Vector Euclidean norm:

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

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

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

Cauchy-Schwartz inequality

$$|\boldsymbol{a}^H \boldsymbol{b}| \leq \|\boldsymbol{a}\| \cdot \|\boldsymbol{b}\|,$$

where equality holds only iff $\boldsymbol{a}$ and $\boldsymbol{b}$ are colinear (i.e. $\boldsymbol{a} = \alpha \boldsymbol{b}$).

Orthogonal vectors:

$$\boldsymbol{a}^H \boldsymbol{b} = \boldsymbol{b}^H \boldsymbol{a} = 0.$$

Example: Consider the output of a linear time-invariant (LTI) system (filter):

$$y(n) = \sum_{k=0}^{N-1} h(k)x(n-k) = \boldsymbol{h}^T \boldsymbol{x}(n)$$

where

$$\boldsymbol{h} = \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(N-1) \end{bmatrix}, \quad \boldsymbol{x}(n) = \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-N+1) \end{bmatrix}.$$

2.3.2 Linear Independence, Vector Spaces, and Basis Vectors

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

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots \alpha_n \boldsymbol{x}_n = 0 \tag{1}$$

implies that $\alpha_i = 0$ for all i. If a set of nonzero α_i can be found so that (1) holds, the vectors are *linearly dependent*. For example, for nonzero α_1,

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

Example 2.3.2:

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{linearly independent}$$

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ \boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{linearly dependent}$$

because $\boldsymbol{x}_1 = \boldsymbol{x}_2 + 2\boldsymbol{x}_3$.

Given N vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N$, consider the set of all vectors $\mathcal{V}$ that may be formed as a linear combination of the vectors $\boldsymbol{v}_i$:

$$\boldsymbol{v} = \sum_{i=1}^{N} \alpha_i \boldsymbol{v}_i.$$

This set forms a *vector space* and the vectors $\boldsymbol{v}_i$ are said to *span* the space $\mathcal{V}$.

If $\boldsymbol{v}_i$'s are linearly independent, they are said to form a *basis* for the space $\mathcal{V}$ and the number of basis vectors N is referred to as the space *dimension*. The basis for a vector space is not unique!

2.3.3 Matrices

$n \times m$ matrix:

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

If $n = m$, then A is a *square* matrix. *Symmetric* matrix:

$$A^T = A.$$

Hermitian matrix:

$$A^H = A.$$

Some properties [apply to transpose ^{T} as well]:

$$(A + B)^H = A^H + B^H, \quad (A^H)^H = A, \quad (AB)^H = B^H A^H.$$

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

$$A = [\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_m] = \begin{bmatrix} \boldsymbol{r}_1^H \\ \boldsymbol{r}_2^H \\ \vdots \\ \boldsymbol{r}_n^H \end{bmatrix} \tag{2}$$

2.3.4 Matrix Inverse

Rank Discussion: The *rank* of $A \equiv \#$ of linearly independent columns in (2) $\equiv \#$ of linearly independent row vectors in (2) [i.e. $\#$ of linearly independent vectors in $\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n\}$]. Important property:

$$\mathrm{rank}(A) = \mathrm{rank}(AA^H) = \mathrm{rank}(A^H A).$$

For any $n \times m$ matrix A: $\mathrm{rank}(A) \le \min\{n, m\}$.

If $\mathrm{rank}(A) = \min\{n, m\}$, A is said to be of *full rank*.

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

$$A^{-1}A = AA^{-1} = I$$

where

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

is called the *identity matrix*.

If a square matrix A (of size $n \times n$) is not of full rank [i.e. $\mathrm{rank}(A) < n$], then it is said to be *singular*.

Properties: $(AB)^{-1} = B^{-1}A^{-1}$, $(A^H)^{-1} = (A^{-1})^H$.

Useful Matrix Inversion Identities

Matrix Inversion Lemma:

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

for *arbitrary square nonsingular* A and C. In the special case where C is a scalar, $B = \boldsymbol{b}$ is a column vector and $D = \boldsymbol{d}^H$ is a row vector. For $C = 1$, we obtain the Woodbury identity:

$$(A + \boldsymbol{b}\boldsymbol{d}^H)^{-1} = A^{-1} - \frac{A^{-1}\boldsymbol{b}\,\boldsymbol{d}^H A^{-1}}{1 + \boldsymbol{d}^H A^{-1}\boldsymbol{b}}.$$

2.3.5 The Determinant and the Trace

The determinant of an $n \times n$ matrix (for any i):

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

where A_{ij} is the $(n-1) \times (n-1)$ matrix formed by deleting the ith row and the jth column of A.

Example:

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

$$\begin{aligned} \det(A) &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \\ &\quad + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}. \end{aligned}$$

Property: an $n \times n$ matrix A is invertible (nonsingular) iff its determinant is nonzero,

$$\det(A) \neq 0.$$

Properties of the determinant (for $n \times n$ matrix A):

$$\det(AB) = \det(A)\det(B), \qquad \det(\alpha A) = \alpha^n \det A$$
$$\det(A^{-1}) = \frac{1}{\det A}, \qquad \det(A^T) = \det A.$$

Another important function of a matrix is *trace*:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

2.3.6 Linear Equations

Many practical DSP problems [e.g. signal modeling, Wiener filtering, etc.] require solving 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:

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{3}$$

where

- A is an $n \times m$ matrix with entries a_{ij},
- $\boldsymbol{x}$ is an m-dimensional vector of x_i's, and
- $\boldsymbol{b}$ is an n-dimensional vector of b_i's.

For $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m]$, we can view (3) as an expansion of $\boldsymbol{b}$:

$$\boldsymbol{b} = \sum_{i=1}^{m} x_i \boldsymbol{a}_i.$$

Square matrix $A : m = n$**.** The nature of the solution depends upon whether or not A is singular.

Nonsingular case:

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}.$$

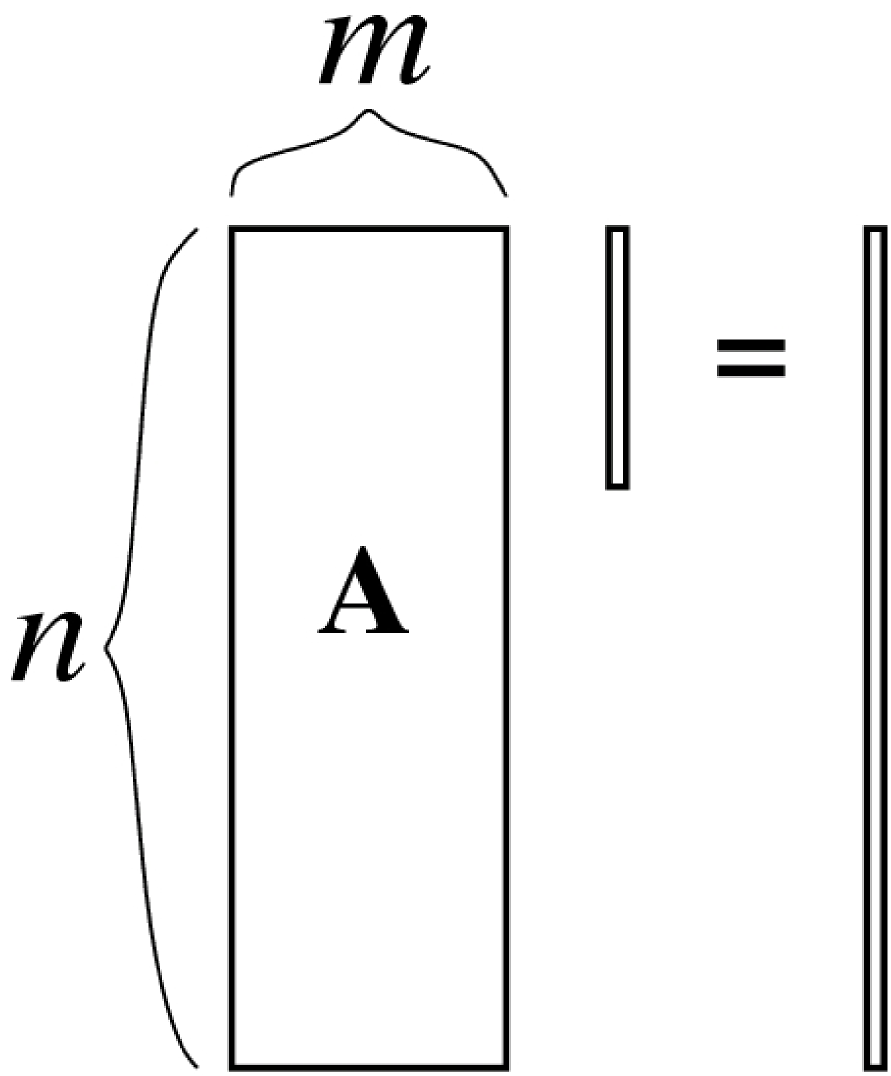
Singular case: no solution or *many solutions*.

Example:

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

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

Rectangular matrix A : $m < n$**.** More equations than unknowns and, in general, *no solution exists*. The system is called *overdetermined*.



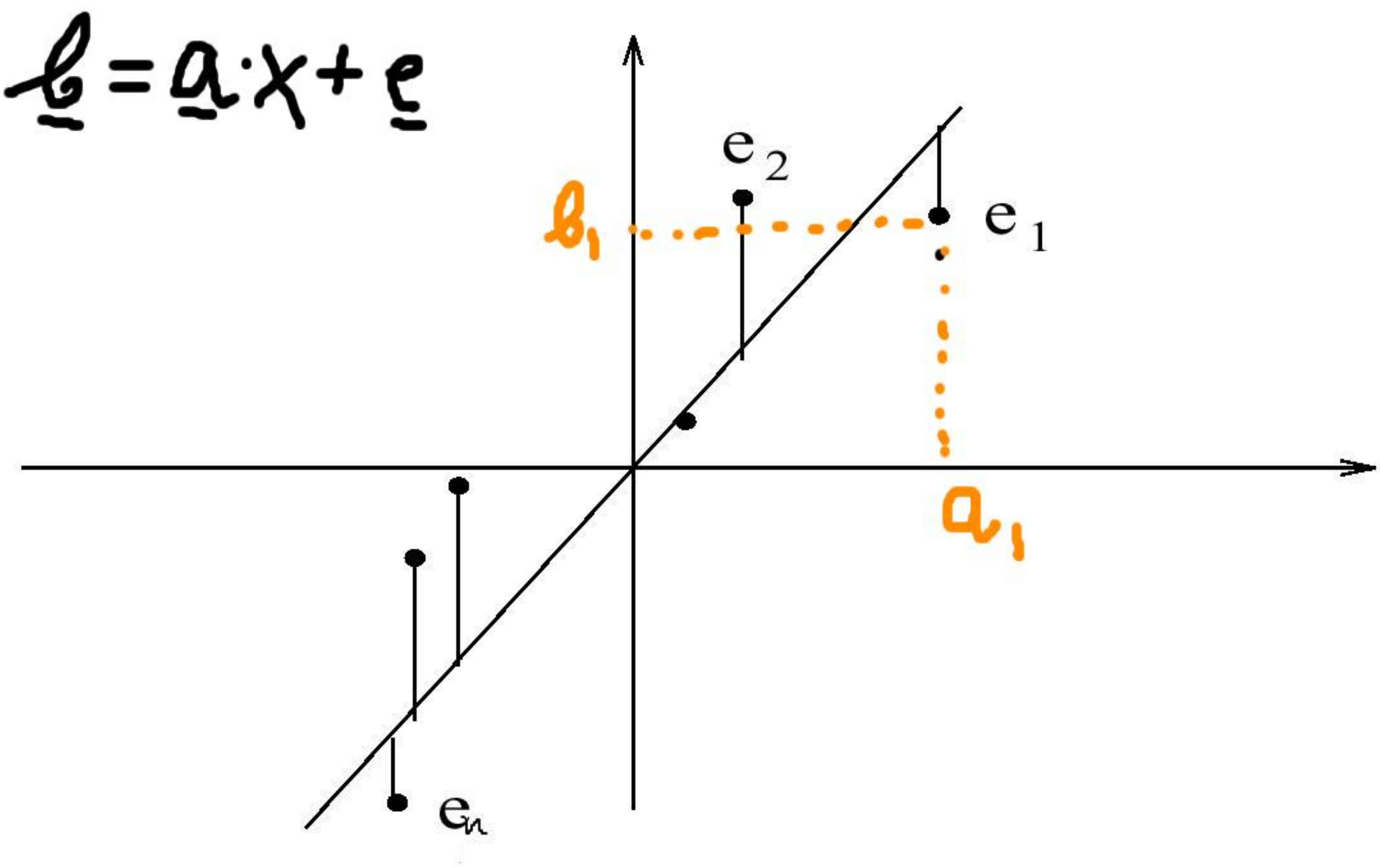
When A is a full-rank matrix and, therefore, $A^H A$ is nonsingular, a common approach is to find the *least-squares solution*, i.e. the vector x that minimizes the Euclidean norm of the error vector:

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

$$
\begin{aligned}
&= (\boldsymbol{b} - A\boldsymbol{x})^H(\boldsymbol{b} - A\boldsymbol{x}) \\
&= \boldsymbol{b}^H\boldsymbol{b} - \boldsymbol{x}^H A^H \boldsymbol{b} - \boldsymbol{b}^H A\boldsymbol{x} + \boldsymbol{x}^H A^H A\boldsymbol{x} \\
&= \left[\boldsymbol{x} - (A^H A)^{-1}A^H\boldsymbol{b}\right]^H A^H A \left[\boldsymbol{x} - (A^H A)^{-1}A^H\boldsymbol{b}\right] \\
&\quad + [\boldsymbol{b}^H\boldsymbol{b} - \boldsymbol{b}^H A(A^H A)^{-1}A^H\boldsymbol{b}].
\end{aligned}
$$

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

$$\boldsymbol{x}_{\text{LS}} = (A^H A)^{-1}A^H\boldsymbol{b}.$$



The best (LS) approximation $\widehat{\boldsymbol{b}}$ to $\boldsymbol{b}$ is given by

$$\widehat{\boldsymbol{b}} = A\boldsymbol{x}_{\mathrm{LS}} = A(A^H A)^{-1} A^H \boldsymbol{b} = P_{_A}\boldsymbol{b},$$

where

$$P_{_A} = A(A^H A)^{-1} A^H$$

is called the *projection matrix* with properties:

$$P_{_A}\boldsymbol{a} = \boldsymbol{a}$$

if the vector $\boldsymbol{a}$ belongs to the column space of A, and

$$P_{_A}\boldsymbol{a} = \boldsymbol{0}$$

if $\boldsymbol{a}$ is orthogonal to the column space of A.

The minimum LS error:

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

where $P_{_A}^{\perp} = I - P_{_A}$ is the projection matrix onto the subspace orthogonal to the column space of A.

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

$$A^H A\boldsymbol{x} = A^H \boldsymbol{b}.$$

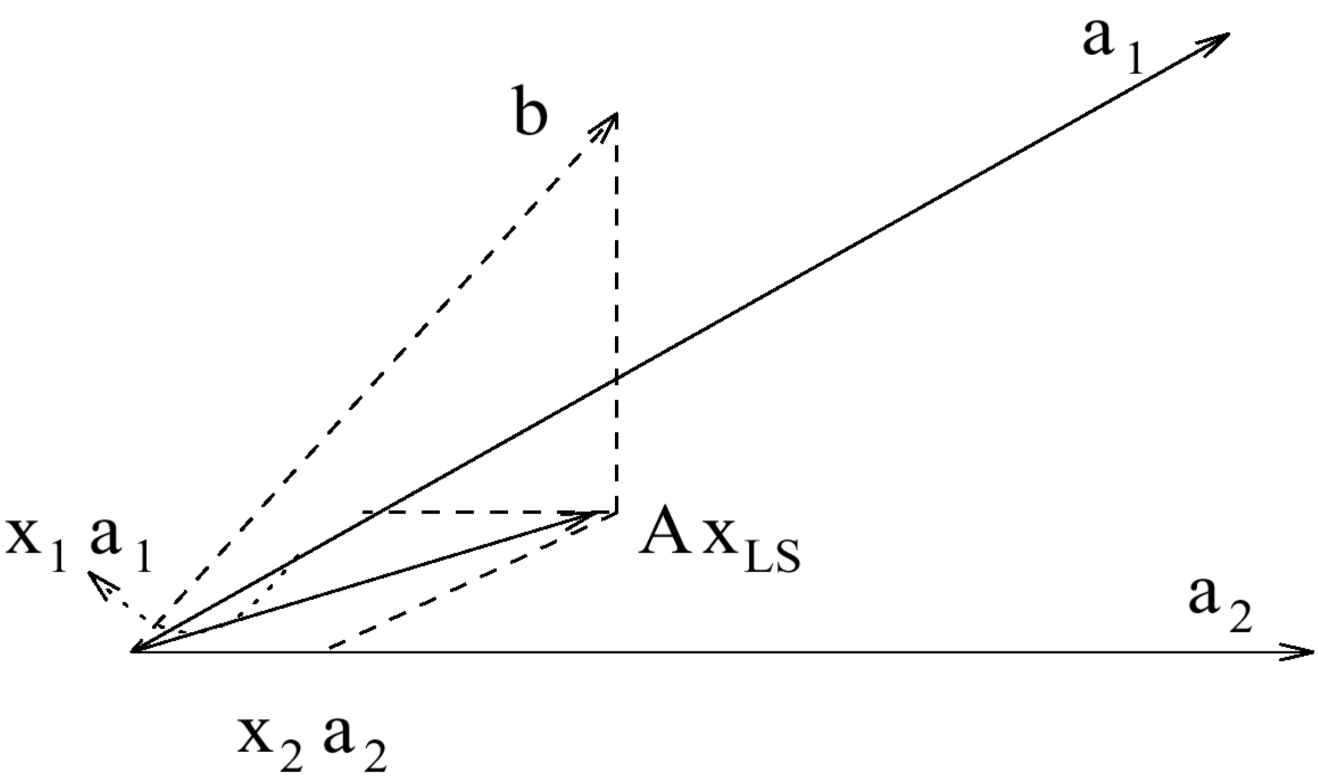
which follow from the *orthogonality principle*:

$$A^H \boldsymbol{e} = \boldsymbol{0}.$$

Illustration of LS Solutions

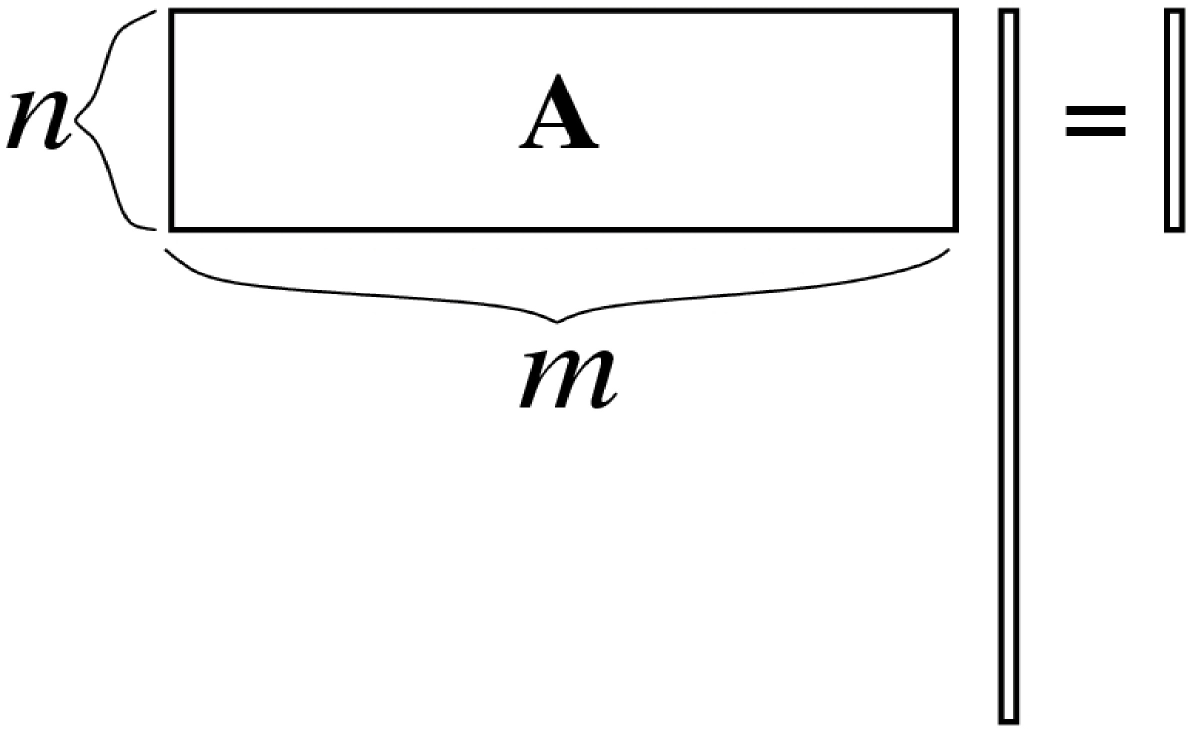
Consider

$$\begin{aligned} A &= [\boldsymbol{a}_1, \boldsymbol{a}_2], \\ \boldsymbol{x}_{\mathrm{LS}} &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \end{aligned}$$



2.3.6 Linear Equations

Rectangular matrix $A : n < m$**.** Fewer equations than unknowns and, provided the equations are consistent, there are *many solutions*. The system is called *underdetermined*.



2.3.7 Special Matrix Forms

Diagonal (square) matrix:

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

Exchange matrix:

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

Toeplitz matrix:

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

Example:

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

2.3.8 Quadratic and Hermitian Forms

Quadratic form of a real symmetric square matrix A:

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

Similarly, *Hermitiam form* of a Hermitian square matrix A:

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

Symmetric (Hermitian) matrices are *positive semidefinite* if $Q(\boldsymbol{x}) \geq 0$ for all nonzero $\boldsymbol{x}$. If $Q(\boldsymbol{x}) > 0$ for all nonzero $\boldsymbol{x}$, then A is said to be *positive definite*.

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

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

2.3.9 Eigenvalues and Eigenvectors

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

$$A\boldsymbol{u} = \lambda \boldsymbol{u},$$

which is equivalent to the following set of homogeneous linear equations:

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

For a nontrivial solution, $A - \lambda I$ needs to be singular. Hence,

$$p(\lambda) = \det(A - \lambda I) = 0.$$

$p(\lambda)$ is called the characteristic polynomial of A, and the n roots, $\lambda_i, i = 1, \ldots, n, \equiv$ the *eigenvalues* of A.

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

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i.$$

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

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

Property 1: The eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ corresponding to *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are linearly independent.

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

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

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

$$\boldsymbol{u}_i^H A \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i^H \boldsymbol{u}_i.$$

Applyng ^{H} to the above equation, we get

$$\boldsymbol{u}_i^H A^H \boldsymbol{u}_i = \lambda_i^* \boldsymbol{u}_i^H \boldsymbol{u}_i.$$

Since A is Hermitian ($A = A^H$), we have

$$\lambda_i^* \boldsymbol{u}_i^H \boldsymbol{u}_i = \boldsymbol{u}_i^H A^H \boldsymbol{u}_i \overbrace{=}^{A = A^H} \boldsymbol{u}_i^H A \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i^H \boldsymbol{u}_i.$$

Thus, $\lambda_i = \lambda_i^*$, i.e. λ_i must be *real*. $\square$

Property 4: A Hermitian matrix is *positive definite* $(A > 0)$ iff the eigenvalues of A are positive.

Similar property holds for negative definite and positive (negative) semi-definite matrices.

A useful relationship between matrix determinant and eigenvalues:

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

Therefore, a matrix is nonsingular (invertible) iff *all* of its eigenvalues are nonzero.

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

Proof. Let λ_i and λ_k be two distinct eigenvalues of A. Then

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i, \quad A\boldsymbol{u}_k = \lambda_k \boldsymbol{u}_k.$$

Multiplying the above equations by $\boldsymbol{u}_k^H$ and $\boldsymbol{u}_i^H$, respectively, we get

$$\boldsymbol{u}_k^H A \boldsymbol{u}_i = \lambda_i \boldsymbol{u}_k^H \boldsymbol{u}_i, \quad \boldsymbol{u}_i^H A \boldsymbol{u}_k = \lambda_k \boldsymbol{u}_i^H \boldsymbol{u}_k.$$

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

$$\boldsymbol{u}_k^H A \boldsymbol{u}_i = \lambda_k \boldsymbol{u}_k^H \boldsymbol{u}_i,$$

leading to

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

Since $\lambda_i \neq \lambda_k$, we have

$$\boldsymbol{u}_k^H \boldsymbol{u}_i = 0.$$

$\square$

Remark: Although verified above only for the case of distinct eigenvalues, it is also true that, for any $n \times n$ Hermitian matrix, there exists a set of n orthonormal eigenvectors.

For any $n \times n$ matrix A having *a set of linearly independent eigenvectors*, we may perform an eigenvalue decomposition (EVD):

$$A = U \Lambda U^{-1}$$

by rewriting the set of equations

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

in the form

$$A[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n] = [\lambda_1 \boldsymbol{u}_1, \lambda_2 \boldsymbol{u}_2, \cdots, \lambda_n \boldsymbol{u}_n],$$

or, equivalently

$$AU = U\Lambda, \tag{4}$$

where

$$\begin{aligned} U &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n], \\ \Lambda &= \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}. \end{aligned}$$

Since we have assumed that *U is nonsingular*, we can right-multiply the equation (4) by U^{-1}.

2.3.9 Eigenvalues and Eigenvectors (Hermitian matrix)

For a Hermitian matrix, we can *always* find an orthonormal set of eigenvectors:

$$U^H U = I.$$

Hence, U is *unitary* (i.e. $U^H = U^{-1}$), and the EVD becomes

$$A = U \Lambda U^H$$

or, equivalently,

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

This result is known as the *spectral theorem*.

Using the unitary property of U, it is easy to find the inverse of a nonsigular Hermitian matrix via EVD:

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

or, equivalently,

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

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

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

$$A = B + \alpha I.$$

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

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

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

$$B\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i.$$

2.3.9 Eigenvalues and Eigenvectors — Trace

We can write the trace of A in terms of its eigenvalues:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i.$$

Similarly,

$$\mathrm{tr}(A^{-1}) = \sum_{i=1}^{n} \frac{1}{\lambda_i}.$$

This property can be easily shown using the EVD and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Denoting the maximum eigenvalue of A by λ_{MAX}, if A is positive semi-definite Hermitian, then

$$\lambda_{\mathrm{MAX}} \leq \sum_{i=1}^{n} \lambda_i = \mathrm{tr}(A).$$

Singular Value Decomposition (SVD)

For a rectangular $n \times m$ matrix A, we may perform the SVD instead of EVD:

$$A = U \Lambda V^H$$

where $UU^H = U^H U = I$ and $VV^H = V^H V = I$ and

$$\begin{aligned} \Lambda &= \begin{cases} [\Lambda(n), 0], & n < m \\ \begin{bmatrix} \Lambda(m) \\ 0 \end{bmatrix}, & n > m \end{cases}, \\ \Lambda(m) &= \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_m\}, \end{aligned}$$

and λ_i's are non-negative. Equivalently,

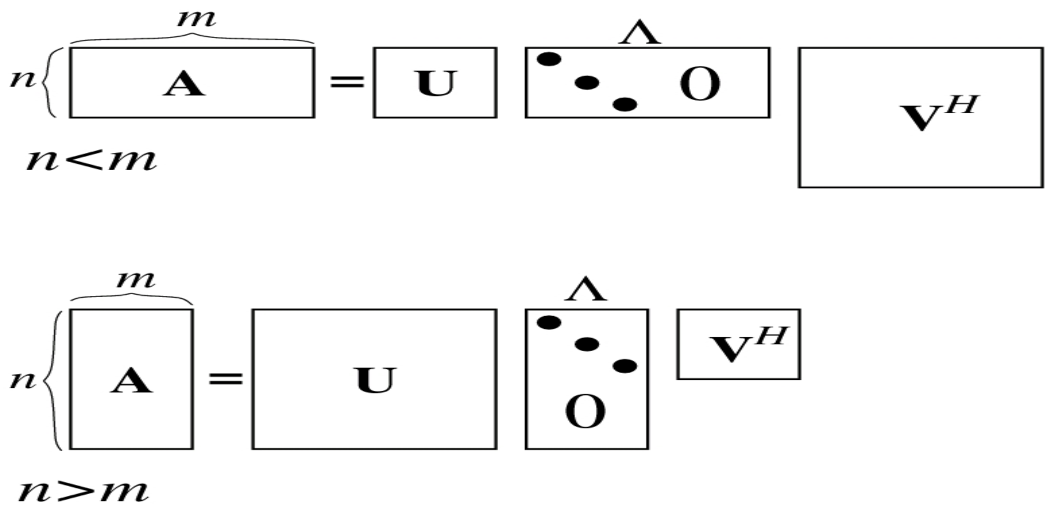
$$A = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^H \quad \text{if } n < m$$

or

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

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

Pictorial representation of the SVD:



Computational Aspects of LS

Solving Normal Equations:

$$A^H A \boldsymbol{x}_{\mathrm{LS}} = A^H \boldsymbol{b}.$$

Define

$$C = A^H A, \quad \boldsymbol{g} = A^H \boldsymbol{b}.$$

Solve

$$C \boldsymbol{x}_{\mathrm{LS}} = \boldsymbol{g},$$

where C is a positive definite Hermitian matrix.

Cholesky Decomposition

Also known as the LDL^H decomposition:

$$C = LDL^H,$$

where

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix} \quad \text{(lower triangular matrix)}$$

and $D = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}, d_i > 0$.

Back-substitution to solve:

$$LDL^H \boldsymbol{x}_{\mathrm{LS}} = \boldsymbol{g}.$$

Cholesky Decomposition Approach to Solving LS: Define $\boldsymbol{y} = DL^H \boldsymbol{x}_{\mathrm{LS}}$. Then

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix}.$$

$$
\begin{aligned}
y_1 &= g_1, \\
y_2 &= g_2 - l_{21} y_1, \\
y_k &= g_k - \sum_{i=1}^{k-1} l_{ki} y_i, \quad k = 1, 2, \dots, n.
\end{aligned}
$$

$$
\begin{bmatrix} 1 & l_{21}^* & \cdots & l_{n1}^* \\ 0 & 1 & \cdots & l_{n2}^* \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = L^H \boldsymbol{x}_{\mathrm{LS}} = D^{-1} \boldsymbol{y} = \begin{bmatrix} \frac{y_1}{d_1} \\ \vdots \\ \frac{y_n}{d_n} \end{bmatrix}.
$$

$$
\begin{aligned}
x_n &= \frac{y_n}{d_n}, \\
x_k &= \frac{y_k}{d_k} - \sum_{i=k+1}^{n} l_{ik}^* x_i, \quad k = n-1, \dots
\end{aligned}
$$

Note: Solving normal equations using this approach may be sensitive to numerical errors.

QR Decomposition Approach to Solving LS

A can be factored as

$$A = QR = [Q_1 \; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1,$$

where Q is unitary (i.e. $QQ^H = Q^H Q$) and R_1 is square upper triangular (Matlab: `qr`). Then

$$(A^H A)^{-1} A^H = R_1^{-1} Q_1^H$$

and $\boldsymbol{x}_{\mathrm{LS}}$ is obtained by solving the following triangular system:

$$R_1 \boldsymbol{x}_{\mathrm{LS}} = Q_1^H \boldsymbol{b}.$$

Note: Numerically more robust than Cholesky. For a large number of overdetermined equations, the QR method needs about $2\times$ more computations compared with Cholesky.