MP-208
Optimal Filtering with Aerospace Applications
Section 2.1: Linear Algebra and Matrices

Prof. Dr. Davi Antônio dos Santos
Instituto Tecnológico de Aeronáutica
www.professordavisantos.com

São José dos Campos - SP
2023

Real Matrix: Definition and Notation

Definition:

A real matrix with dimensions $n \times m$ is the following bi-dimensional arrangement:

$$\mathbf{A} \triangleq \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \in \mathbb{R}^{n \times m}, \tag{1}$$

in which the elements $a_{ij} \in \mathbb{R}$, $\forall i \in \{1, 2, ..., n\}$ and $\forall j \in \{1, 2, ..., m\}$.

Notation:

- Matrices: boldface upper-case letters.
- Transpose matrix: $\mathbf{A}^{\mathrm{T}}$.
- Scalars and subscripts/superscripts: italic lower-case letters.

Real Vector: Definition and Notation

Definition:

A real vector with dimension n is a real matrix (column) of dimensions $n \times 1$:

$$\mathbf{a} \triangleq \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n \tag{2}$$

in which $a_i \in \mathbb{R}$, $i \in \{1, ..., n\}$, are its components.

Notation:

- Vectors: boldface lower-case letters.

Diagonal Matrix

Definition:

The square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a diagonal matrix if all its elemets out of the primary diagonal are zeros:

$$\mathbf{A} \triangleq \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \tag{3}$$

Remarks:

- Frequent notation: $\mathbf{A} = \mathrm{diag}(a_{11}, ..., a_{nn})$.
- Block-diagonal matrix: $\mathbf{B} = \mathrm{diag}(\mathbf{A}_1, ..., \mathbf{A}_k)$, where $\mathbf{A}_i$, $i = 1, ..., k$, are matrices.
- Identity matrix: If $a_{11} = a_{22} = \ldots = a_{nn} = 1$, then $\mathbf{A} \triangleq \mathbf{I}_n$ is called the identity matrix (of dimension $n \times n$).

Symmetric and Skew-Symmetric Matrices

Symmetric Matrix:

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric if

$$\mathbf{A} = \mathbf{A}^{\mathrm{T}} \tag{4}$$

In other words, the elements of $\mathbf{A}$ are such that $a_{ij} = a_{ji}, \forall i, j = 1, ..., n$.

Skew-Symmetric Matrix:

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ if skew-symmetric if

$$\mathbf{A} = -\mathbf{A}^{\mathrm{T}} \tag{5}$$

In other words, the elements of $\mathbf{A}$ are such that $a_{ii} = 0, \forall i$ and $a_{ij} = -a_{ji}, \forall i \neq j$.

Symmetric and Skew-Symmetric Matrices

Interesting result:

Any square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ can be written as the sum of a symmetric and a skew-symmetric matrix:

$$\mathbf{A} = \frac{\mathbf{A} + \mathbf{A}^{\mathrm{T}}}{2} + \frac{\mathbf{A} - \mathbf{A}^{\mathrm{T}}}{2} \tag{6}$$

This is the so-called Toeplitz decomposition.

Lower and Upper Triangular Matrices

Lower Triangular Matrix:

It is a square matrix with zero elements over the primary diagonal:

$$\mathbf{L} \triangleq \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{7}$$

Upper Triangular Matrix:

It is a square matrix with zero elements under the primary diagonal:

$$\mathbf{U} \triangleq \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \tag{8}$$

Trace of a Matrix

Definition:

The trace of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the sum of the elements of its primary diagonal:

$$\mathrm{tr}\left(\mathbf{A}\right) \triangleq \sum_{i=1}^{n} a_{ii} \tag{9}$$

Interesting Results (A,B,C are square and α is a scalar):

- $\mathrm{tr}\left(\mathbf{A}\right) = \mathrm{tr}\left(\mathbf{A}^{\mathrm{T}}\right)$
- $\mathrm{tr}\left(\mathbf{ABC}\right) = \mathrm{tr}\left(\mathbf{CAB}\right) = \mathrm{tr}\left(\mathbf{BCA}\right)$
- $\mathrm{tr}\left(\mathbf{AB}\right) = \mathrm{tr}\left(\mathbf{BA}\right)$
- $\mathrm{tr}\left(\mathbf{A} + \mathbf{B}\right) = \mathrm{tr}\left(\mathbf{A}\right) + \mathrm{tr}\left(\mathbf{B}\right)$
- $\mathrm{tr}\left(\alpha\mathbf{A}\right) = \alpha\,\mathrm{tr}\left(\mathbf{A}\right)$

Determinant of a Matrix

The determinant of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a kind of signed volume and can be computed recursively by the Laplace Formula:

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}) \quad (10)$$

where $\mathbf{A}_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ is a sub-matrix obtained from $\mathbf{A}$ by excluding its ith row and jth column.

Interesting Results:

- $\det(\mathbf{A}) = \det(\mathbf{A}^{\mathrm{T}})$.
- $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.
- $\det(\alpha\mathbf{A}) = \alpha^{n}\det(\mathbf{A})$.

Rank of a Matrix

Definition:

The rank of a matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ is its number of linearly independent rows or columns. We denote the rank of $\mathbf{A}$ by:

$$\operatorname{rank}(\mathbf{A}) \tag{11}$$

Remarks:

- Consider a matrix $\mathbf{A} \in \mathbb{A}^{n \times m}$. In this case, if $\operatorname{rank}(\mathbf{A}) = \min(n, m)$, then we say that $\mathbf{A}$ has full rank.
- If $\mathbf{A} \in \mathbb{R}^{n \times m}$ is not full-rank, it is called a rank-deficient matrix.
- $\operatorname{rank}(\mathbf{A}) = \dim\big(R(\mathbf{A})\big)$, where $R(\mathbf{A})$ is the column space of $\mathbf{A}$.

Cofactor and Adjoint Matrices

Cofactor:

The cofator $\tilde{a}_{ij} \in \mathbb{R}$ relative to the element a_{ij} of $\mathbf{A} \in \mathbb{R}^{n \times n}$ is given by

$$\tilde{a}_{ij} \triangleq (-1)^{i+j} \det \left(\mathbf{A}_{ij} \right) \tag{12}$$

Cofactor matrix:

The cofactor matrix relative to $\mathbf{A} \in \mathbb{R}^{n \times n}$ is

$$\mathrm{cof}(\mathbf{A}) \triangleq \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{a}_{22} & \dots & \tilde{a}_{2n} \\ \vdots & \vdots & & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \dots & \tilde{a}_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{13}$$

Adjoint Matrix:

The adjoint matrix relative to $\mathbf{A} \in \mathbb{R}^{n \times n}$ is $\mathrm{adj}(\mathbf{A}) \triangleq \mathrm{cof}(\mathbf{A})^{\mathrm{T}}$.

Singular and Inverse Matrices

Singular Matrix:

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is said to be singular if $\det(\mathbf{A}) = 0$. Otherwise, $\mathbf{A}$ is said to be nonsingular.

Inverse Matrix:

Consider a nonsingular square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. The inverse of $\mathbf{A}$, which we denote by $\mathbf{A}^{-1} \in \mathbb{R}^{n \times n}$, is such that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_n. \tag{14}$$

Interesting Result:

- $\mathbf{A}^{-1} = \dfrac{\mathrm{adj}(\mathbf{A})}{\det(\mathbf{A})}$.
- $\det\left(\mathbf{A}^{-1}\right) = 1/\det(\mathbf{A})$.

Inverse of a Partitioned Matrix

Interesting Result:

Consider the matrices $\mathbf{P}_{11} \in \mathbb{R}^{n_1 \times n_1}$, $\mathbf{P}_{12} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{P}_{21} \in \mathbb{R}^{n_2 \times n_1}$ and $\mathbf{P}_{22} \in \mathbb{R}^{n_2 \times n_2}$. One can show that

$$\begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \tag{15}$$

where

$$\mathbf{V}_{11} = \left(\mathbf{P}_{11} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\right)^{-1} \tag{16}$$

$$\mathbf{V}_{12} = -\mathbf{V}_{11}\mathbf{P}_{12}\mathbf{P}_{22}^{-1} \tag{17}$$

$$\mathbf{V}_{21} = -\mathbf{P}_{22}^{-1}\mathbf{P}_{21}\mathbf{V}_{11} \tag{18}$$

$$\mathbf{V}_{22} = \left(\mathbf{P}_{22} - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\right)^{-1} \tag{19}$$

Matrix Inversion Lemma

Another Interesting Result:

Consider the matrices **P**, **R**, and **H** with appropriate dimensions. Assume that **P** and **R** are nonsingular. One can show that

$$\left(\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\right)^{-1} = \mathbf{P} - \mathbf{P}\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P} \qquad (20)$$

Remark:

In future chapters, the above result is used to derive the recursive least squares and the information filter algorithms.

Jacobian Matrix

Definition:

Consider a differentiable vectorial function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Denote its independent variable by $\mathbf{x} \in \mathbb{R}^n$. The Jacobian matrix of $\mathbf{f}$ is given by

$$\frac{d\mathbf{f}}{d\mathbf{x}} \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{m \times n} \tag{21}$$

where f_i shortly denotes the ith component $f_i(\mathbf{x}) \in \mathbb{R}$ of the vector $\mathbf{f}(\mathbf{x})$, while x_j denotes the jth component of $\mathbf{x}$.

Hessian Matrix

Definition:

Consider a twice differentiable scalar function $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Denote its independent variable by $\mathbf{x} \in \mathbb{R}^n$. The Hessian matrix of g is the symmetric matrix given by

$$\frac{d^2 g}{d\mathbf{x}^2} \triangleq \begin{bmatrix} \frac{\partial^2 g}{\partial x_1^2} & \frac{\partial^2 g}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 g}{\partial x_1 \partial x_n} \\ \frac{\partial^2 g}{\partial x_2 \partial x_1} & \frac{\partial^2 g}{\partial x_2^2} & \cdots & \frac{\partial^2 g}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 g}{\partial x_n \partial x_1} & \frac{\partial^2 g}{\partial x_n \partial x_2} & & \frac{\partial^2 g}{\partial x_n^2} \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{22}$$

where g is an abbreviation for $g(\mathbf{x})$.

Eigenvalues and Eigenvectors

Definition:

Consider a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. The eigenvalues $\lambda_i \in \mathbb{C}$ and eigenvectors $\boldsymbol{\nu}_i \in \mathbb{R}^n$, $i = 1, ..., n$, relative to $\mathbf{A}$ are such that

$$\mathbf{A}\boldsymbol{\nu}_i = \lambda_i \boldsymbol{\nu}_i, \ \forall i \tag{23}$$

Remark:

- If the eigenvalues λ_i are all distinct from one another, then the corresponding eigenvectors $\boldsymbol{\nu}_i$ are all linearly independent.
- An eigenvector is parallel to the vector resulting from its premultiplication with $\mathbf{A}$.
- The eigenvalues are obtained by solving the characteristic (polynomial) equation $\det(\lambda \mathbf{I}_n - \mathbf{A}) = 0$.

Matrix Square Root: Cholesky Decomposition

Result:

Consider a real symmetric positive-definite [1] matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. One can show that there is a unique decomposition of $\mathbf{A}$ into the form:

$$\mathbf{A} = \mathbf{L}\mathbf{L}^{\mathrm{T}} \tag{24}$$

where $\mathbf{L}$ is a lower-triangular matrix with positive diagonal elements.

Remarks:

- If $\mathbf{A}$ is symmetric and indefinite, one can use the LDL^{T} decomposition.
- It will be used to compute the square root of a covariance matrix in the UKF algorithm.

[1] $\mathbf{A}$ is said to be PD if $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0, \forall \mathbf{x} \in \mathbb{R}^n / \{0\}$. Matrix $\mathbf{A}$ is PD iff all its eigenvalues are positive.

Matrix Exponential

Definition:

The exponential of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is defined by

$$\exp(\mathbf{A}) \triangleq \mathbf{I}_n + \mathbf{A} + \frac{1}{2!}\mathbf{A}^2 + \ldots + \frac{1}{k!}\mathbf{A}^k + \ldots \tag{25}$$

Remark:

There are many ways to compute (in general, approximately) the exponential of a matrix. For this course, I suggest [2]:

- The Sylvester method
- The diagonalization method

[2] For more possibilities, see (Moler and Van Loan, 2003).

Inner Product and Norm of Vectors

Inner Product:

Consider two real vectors with the same dimension $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^n$. The inner product of $\mathbf{a}$ with $\mathbf{b}$ is denoted by $\langle \mathbf{a}, \mathbf{b} \rangle \in \mathbb{R}$ and defined by

$$\langle \mathbf{a}, \mathbf{b} \rangle \triangleq \mathbf{a}^{\mathrm{T}} \mathbf{b} \tag{26}$$

l_2-Norm:

The l_2-norm of a vector $\mathbf{a} \in \mathbb{R}^n$ is denoted by $\|\mathbf{a}\| \in \mathbb{R}$ and defined by

$$\|\mathbf{a}\| \triangleq \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} \tag{27}$$

Remarks:

- $\|\mathbf{a}\| > 0, \ \forall \mathbf{a} \neq \mathbf{0}_{n \times 1}$ and $\|\mathbf{a}\| = 0$ only if $\mathbf{a} = \mathbf{0}_{n \times 1}$.
- $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle$, etc.

Norm of a Matrix

Definition: Frobenius Norm

The Frobenius norm of a matrix $\mathbf{A} \in \mathbb{R}^{n \times p}$ is defined by

$$\|\mathbf{A}\|_F \triangleq \sqrt{\mathrm{tr}\left(\mathbf{A}\mathbf{A}^{\mathrm{T}}\right)} \tag{28}$$

Definition: l_2-Norm

The norm of $\mathbf{A} \in \mathbb{R}^{n \times p}$ induced by the l_2-norm of a vetor $\mathbf{a} \in \mathbb{R}^p$ is defined by

$$\|\mathbf{A}\|_2 \triangleq \max_{\mathbf{a} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{a}\|}{\|\mathbf{a}\|} \tag{29}$$

Condition Number

Definition:

Consider a symmetric positive-definite square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. The condition number of $\mathbf{A}$ is

$$\kappa(\mathbf{A}) \triangleq \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \quad (30)$$

where $\|.\|$ can be any matrix norm.

Remark:

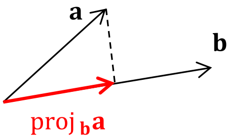
Large values of $\kappa(\mathbf{A})$ indicate that $\mathbf{A}$ is ill-conditioned (*i.e.*, it is "almost" singular!).

Orthogonal Projection of a Vector

Definition:

The orthogonal projection of a vector $\mathbf{a} \in \mathbb{R}^n$ on a vector $\mathbf{b} \in \mathbb{R}^n$ is the vector

$$\mathrm{proj}_{\mathbf{b}}\, \mathbf{a} \triangleq \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{b}\|^2} \mathbf{b} \in \mathbb{R}^n \tag{31}$$



Remark:

One can show that

$$(\mathbf{a} - \mathrm{proj}_{\mathbf{b}}\, \mathbf{a}) \perp \mathbf{b} \tag{32}$$

QR Decomposition

Result:

Consider a real nonsingular matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. One can show that there exists a unique decomposition of $\mathbf{A}$ in the form:

$$\mathbf{A} = \mathbf{QR} \tag{33}$$

where $\mathbf{Q}$ is an orthonormal matrix and $\mathbf{R}$ is an upper triangular matrix.

Remark:

- There are many methods to compute the above decomposition. In this course we can adopt the Gram-Schmidt process.
- In general, it can be used to efficiently solve systems of linear equations or to obtain a matrix inverse.
- It can be used to improve numerical properties of filters.

LU Decomposition

Definition:

Consider a real nonsingular matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. Its LU decomposition is given by:

$$\mathbf{A} = \mathbf{L}\mathbf{U} \tag{34}$$

where $\mathbf{L}$ is a lower triangular matrix (with ones in the primary diagonal) and $\mathbf{U}$ is an upper triangular matrix (not necessarily with ones in the primary diagonal).

Remark:

- $\mathbf{U}$ can be obtained by Gauss elimination and $\mathbf{L}$ is formed with the multipliers of the Gauss elimination process (an example is given on the board).
- It can be used to improve numerical properties of filters.

References. . .

References

- Golub, G. H.; Van Loan, C. F. **Matrix Computations**. Johns Hopkins University Press, 1996.
- Bernstein, D. S. **Matrix Mathematics**. Princeton University Press, 2005.
- Moler, C.; Van Loan, C. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. **Siam Review**, 45(1), 2003.