

# Notation

Rubem Vasconcelos Pacelli  
rubem.engenharia@gmail.com

Department of Teleinformatics Engineering,  
Federal University of Ceará.  
Fortaleza, Ceará, Brazil.

Version: August 15, 2023

## Contents

### 1 Font notation

### 2 Signals and functions

2.1	Time indexing . . . . .
2.2	Common signals . . . . .
2.3	Common functions . . . . .
2.4	Operations and symbols . . . . .
2.5	Digital signal processing . . . . .
2.6	Transformations . . . . .

### 3 Probability, statistics, and stochastic processes

3.1	Operators and symbols . . . . .
3.2	Stochastic processes . . . . .
3.3	Functions . . . . .
3.4	Distributions . . . . .

### 4 Machine learning, optimization theory, and statistical signal processing

4.1	Matrix Calculus (in denominator layout) . . . . .
4.2	Statistics: estimation and detection theory . . . . .
4.3	Signals, (hyper)parameters, system performance, and criteria . . . . .

### 5 Linear Algebra

5.1	Common matrices and vectors . . . . .
5.2	Indexing . . . . .
5.3	General operations . . . . .
5.4	Operations with matrices and tensors . . . . .
5.5	Operations with vectors . . . . .
5.6	Decompositions . . . . .

5.7	Spaces and sets . . . . .	
5.7.1	Common spaces and sets . . . . .	
5.7.2	Convex sets (or spaces) . . . . .	
5.7.3	Spaces from matrices or vectors . . . . .	
5.8	Set operations . . . . .	
5.9	Inequalities . . . . .	
<b>6</b>	<b>Communication systems</b>	
6.1	Common symbols . . . . .	
6.2	Fading multipath channels . . . . .	
<b>7</b>	<b>Discrete mathematics</b>	
7.1	Quantifiers, inferences . . . . .	
7.2	Propositional Logic . . . . .	
7.3	Operations . . . . .	
<b>8</b>	<b>Vector Calculus</b>	
<b>9</b>	<b>Electromagnetic waves</b>	
<b>10</b>	<b>Generic mathematical symbols</b>	
<b>11</b>	<b>Abbreviations</b>	

## 1 Font notation

$a, b, c, \dots, A, B, C, \dots$	Scalars or tuples (represented in parentheses [41] or in angle brackets [11])
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$	Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$	Matrices
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$	Tensors
$A, B, C, \dots, \mathcal{A}, \mathcal{B}, \mathcal{C}, \dots, \mathbb{A}, \mathbb{B}, \mathbb{C}, \dots$	Sets

## 2 Signals and functions

### 2.1 Time indexing

$x(t)$	Continuous-time $t$
$x[n], x[k], x[m], x[i], \dots$ $x_n, x_k, x_m, x_i, \dots$ $x(n), x(k), x(m), x(i), \dots$	Discrete-time $n, k, m, i, \dots$ (parenthesis should be adopted only if there are no continuous-time signals in the context to avoid ambiguity)
$x[((n-m))_N]$ [33], $x((n-m))_N$ [27]	Circular shift in $m$ samples within a $N$ -samples window

### 2.2 Common signals

$\delta(t)$	Delta function
$\delta[n], \delta_{i,j}$	Kronecker function ( $n = i - j$ )
$h(t), h[n]$	Impulse response (continuous and discrete time)
$\tilde{x}[n], \tilde{x}(t)$	Periodic discrete- or continuous-time signal
$\hat{x}[n], \hat{x}(t)$	Estimate of $x[n]$ or $x(t)$
$\dot{x}[m]$	Interpolation of $x[n]$

### 2.3 Common functions

$\mathcal{O}(\cdot), \mathcal{O}(\cdot)$	Big-O notation
$\Gamma(\cdot)$	Gamma function
$\mathcal{Q}(\cdot)$	Quantization function
$\text{sgn}(\cdot)$	Signum function
$\tanh(\cdot)$	Hyperbolic tangent function

$I_\alpha(\cdot)$	Modified Bessel function of the first kind and order $\alpha$
$\binom{n}{k}$	Binomial coefficient

## 2.4 Operations and symbols

$f : A \rightarrow B$	A function $f$ whose domain is $A$ and codomain is $B$
$\mathbf{f} : A \rightarrow \mathbb{R}^n$	A vector-valued function $\mathbf{f}$ , i.e., $n \geq 2$
$f^n, x^n(t), x^n[k]$	$n$ th power of the function $f$ , $x[n]$ or $x(t)$
$f^{(n)}, x^{(n)}(t)$	$n$ th derivative of the function $f$ or $x(t)$
$f', f^{(1)}, x'(t)$	1th derivative of the function $f$ or $x(t)$
$f'', f^{(2)}, x''(t)$	2th derivative of the function $f$ or $x(t)$
$\arg \max [x \in \mathcal{A}] f(x)$	Value of $x$ that minimizes $x$
$\arg \min [x \in \mathcal{A}] f(x)$	Value of $x$ that minimizes $x$
$f(\mathbf{x}) = \inf_{\mathbf{y} \in \mathcal{A}} g(\mathbf{x}, \mathbf{y})$	Infimum, i.e., $f(\mathbf{x}) = \min \{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in \mathcal{A} \wedge (\mathbf{x}, \mathbf{y}) \in \text{dom}(g)\}$ , which is the greatest lower bound of this set [10]
$f(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{A}} g(\mathbf{x}, \mathbf{y})$	Supremum, i.e., $f(\mathbf{x}) = \max \{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in \mathcal{A} \wedge (\mathbf{x}, \mathbf{y}) \in \text{dom}(g)\}$ , which is the least upper bound of this set [10]
$f \circ g$	Composition of the functions $f$ and $g$
$*$	Convolution (discrete or continuous)
$\otimes$ [18], $\bigcircled{\text{N}}$ [33]	Circular convolution

## 2.5 Digital signal processing

$T_s$ [27], $T$ [33]	Sampling period
$f_s, F_s$ [27]	Sampling frequency (in Hz or samples per second [27, chapter 3]), i.e., $1/T_s$

$f$	Continuous linear frequency (in Hz). Apparently, there is no notation for the discrete linear frequency, we use $\omega$ only. However, in [27], the uppercase letters $F$ and $\Omega$ are used to denote the continuous-time frequency, while the lowercase $f$ and $\omega$ denote the discrete-time frequency (Oppenheim [33] does not do it!)
$\Omega$ [27]	Continuous angular frequency (in rad/s), that is, $2\pi f$ .
$\Omega_s$	Sampling frequency (in rad/s), i.e., $2\pi f_s$
$\omega$	Discrete angular frequency, i.e., $\Omega T_s$ [27, eq (3.27)]. As $\omega$ is also used to denote continuous angular frequency outside the DSP context, it is always convenient to state that it denotes the discrete frequency when it does
$W_N$	Twiddle factor, $e^{-j\frac{2\pi}{N}}$ [27]
$N$	Number of samples in the DFT/FFT
$\mathcal{R}_N[n]$	Rectangular window used to cut off the discrete sequences [27]
$\Omega_N$ [33], $B$	One-sided effective bandwidth of the continuous-time signal spectrum
$\omega_s$ [27]	Stop frequency
$\omega_p$ [27]	Pass frequency
$\Delta\omega$ [27]	$\omega_s - \omega_p$
$\omega_c$ [27]	Cutoff frequency
$s(t)$	Impulse train
$\text{gdr}[H(e^{j\omega})]$ [33]	Group delay of $H(e^{j\omega})$
$\angle H(e^{j\omega})$ [33]	Phase response of $H(e^{j\omega})$
$ H(e^{j\omega}) $ [33]	Magnitude (or gain) of $H(e^{j\omega})$
$x_c(t)$ [33], $x(t)$	Continuous-time signal
$x_s(t)$	Sampled version of $x(t)$ , i.e., $x(t)s(t)$
$x_r(t)$	Reconstruction of $x(t)$ from interpolation
$\tilde{x}[n]$	Periodic extension of the the aperiodic signal $x[n]$

## 2.6 Transformations

$\mathcal{F}\{\cdot\}$ [33, section 2.9]	Fourier transform (FT)
--	------------------------

DTFT $\{\cdot\}$ , DFS $\{\cdot\}$ , FFT $\{\cdot\}$	Discrete-time Fourier Transform (DTFT), Discrete Fourier Transform (DFT), Discrete Fourier Series (DFS), respectively
$\mathcal{L}\{\cdot\}$	Laplace transform
$\mathcal{Z}\{\cdot\}$	z-transform
$\hat{x}(t), \hat{x}[n]$	Hilbert transform of $x(t)$ or $x[n]$
$X(s)$	Laplace transform of $x(t)$
$X(f)$	Fourier transform (FT) (in linear frequency, Hz) of $x(t)$
$X(j\omega)$	Fourier transform (FT) (in angular frequency, rad/sec) of $x(t)$
$X(e^{j\omega})$	Discrete-time Fourier transform (DTFT) of $x[n]$
$X[k], X(k), X_k$	Discrete Fourier transform (DFT) or fast Fourier transform (FFT) of $x[n]$ , or even the Fourier series (FS) of the periodic signal $x(t)$
$\tilde{X}[k], \tilde{X}(k), \tilde{X}_k$	Discrete Fourier series (DFS) of $\tilde{x}[n]$
$X(z)$	z-transform of $x[n]$

### 3 Probability, statistics, and stochastic processes

#### 3.1 Operators and symbols

$\mathbf{E}[\cdot], \mathbf{E}[\cdot] \text{ [32]}, E[\cdot], \mathbb{E}[\cdot] \text{ [17]}$	Statistical expectation operator
$\cdot, \mathbf{E}_u[\cdot] \text{ [32]}, E_u[\cdot], \mathbb{E}_u[\cdot]$	Statistical expectation operator with respect to $u$
$\langle \cdot \rangle$	Ensemble average
$\text{var}[\cdot] \text{ [32]}, \text{VAR}[\cdot] \text{ [9, 26, 31, 35]}$	Variance operator
$\text{var}_u[\cdot][\cdot], \text{VAR}_u[\cdot]$	Variance operator with respect to $u$
$\text{cov}[\cdot], \text{COV}[\cdot]$	Covariance operator [9]
$\text{cov}_u[\cdot], \text{COV}_u[\cdot]$	Covariance operator with respect to $u$
$\mu_x$	Mean of the random variable $x$
$\mathbf{\mu}_x, \mathbf{m}_x$	Mean vector of the random variable $\mathbf{x}$ [12]
$\mu_n$	$n$ th-order moment of a random variable
$\sigma_x^2, \kappa_2$	Variance of the random variable $x$
$\mathcal{K}_x, \mu_4$	Kurtosis (4th-order moment) of the random variable $x$

$\kappa_n$	$n$ th-order cumulant of a random variable
$\rho_{x,y}$	Pearson correlation coefficient between $x$ and $y$
$a \sim P$	Random variable $a$ with distribution $P$
$\mathcal{R}$	Rayleigh's quotient

### 3.2 Stochastic processes

PS: Estimated terms come with a hat above of it

$r_x(\tau)$ [32], $R_x(\tau)$	Autocorrelation function of the signal $x(t)$ or $x[n]$
$S_x(f), S_x(j\omega)$	Power spectral density (PSD) of $x(t)$ in linear ( $f$ ) or angular ( $\omega$ ) frequency
$S_{x,y}(f), S_{x,y}(j\omega)$	Cross PSD of $x(t)$ and $y(t)$ in linear or angular ( $\omega$ ) frequency
$\mathbf{R}_x$	(Auto)correlation matrix of $\mathbf{x}(n)$
$r_{x,d}(\tau), R_{x,d}(\tau)$	Cross-correlation between $x[n]$ and $d[n]$ or $x(t)$ and $d(t)$ [32]
$\mathbf{R}_{xy}$	Cross-correlation matrix of $\mathbf{x}(n)$ and $\mathbf{y}(n)$
$\mathbf{r}_{xd}$ [25], $\mathbf{p}_{xd}$ [17]	Cross-correlation vector between $\mathbf{x}(n)$ and $d(n)$
$c_x(\tau), C_x(\tau)$	Autocovariance function of the signal $x(t)$ or $x[n]$ [32]
$\mathbf{C}_x, \mathbf{K}_x, \mathbf{\Sigma}_x, \text{cov}[\mathbf{x}]$	(Auto)covariance matrix of $\mathbf{x}$ [9, 26, 31, 35, 42]
$\tilde{\mathbf{C}}_x$ [35]	Pseudocovariance matrix of $\mathbf{x}$
$c_{xy}(\tau), C_{xy}(\tau)$	Cross-covariance function of the signal $x(t)$ or $x[n]$ [32]
$\mathbf{C}_{xy}, \mathbf{K}_{xy}, \mathbf{\Sigma}_{xy}$	Cross-covariance matrix of $\mathbf{x}$ and $\mathbf{y}$

### 3.3 Functions

$Q(\cdot)$	$Q$ -function, i.e., $P[\mathcal{N}(0, 1) > x]$ [35]
$\text{erf}(\cdot)$	Error function [35]
$\text{erfc}(\cdot)$	Complementary error function i.e., $\text{erfc}(x) = 2Q(\sqrt{2}x) - \text{erf}(x)$ [35]
$P[A]$	Probability of the event or set $A$ [31]

$p(\cdot), f(\cdot)$	Probability density function (PDF) or probability mass function (PMF) [31]
$p(x   A)$	Conditional PDF or PMF [31]
$F(\cdot)$	Cumulative distribution function (CDF)
$\Phi_x(\omega), M_x(j\omega), E[e^{j\omega x}]$	First characteristic function (CF) of $x$ [35, 41]
$M_x(t), \Phi_x(-jt), E[e^{tx}]$	Moment-generating function (MGF) of $x$ [35, 41]
$\Psi_x(\omega), \ln \Phi_x(\omega), \ln E[e^{j\omega x}]$	Second characteristic function
$K_x(t), \ln E[e^{tx}], \ln M_x(t)$	Cumulant-generating function (CGF) of $x$ [26]

### 3.4 Distributions

$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution of a random variable with mean $\mu$ and variance $\sigma^2$
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian distribution of a random variable with mean $\mu$ and variance $\sigma^2$
$\mathcal{N}(\mathbf{\mu}, \mathbf{\Sigma})$	Gaussian distribution of a vector random variable with mean $\mathbf{\mu}$ and covariance matrix $\mathbf{\Sigma}$
$\mathcal{CN}(\mathbf{\mu}, \mathbf{\Sigma})$	Complex Gaussian distribution of a vector random variable with mean $\mathbf{\mu}$ and covariance matrix $\mathbf{\Sigma}$
$\mathcal{U}(a, b)$	Uniform distribution from $a$ to $b$
$\chi^2(n), \chi_n^2$	Chi-square distribution with $n$ degree of freedom (assuming that the Gaussians are $\mathcal{N}(0, 1)$ )
$\text{Exp}(\lambda)$	Exponential distribution with rate parameter $\lambda$
$\Gamma(\alpha, \beta)$	Gamma distribution with shape parameter $\alpha$ and rate parameter $\beta$
$\Gamma(\alpha, \theta)$	Gamma distribution with shape parameter $\alpha$ and scale parameter $\theta = 1/\beta$
$\text{Nakagami}(m, \Omega)$	Nakagami-m distribution with shape parameter or fading figure $m$ and spread, scale, or shape parameter $\Omega$
$\text{Rayleigh}(\sigma)$	Rayleigh distribution with scale parameter $\sigma$



Rayleigh( $\Omega$ )	Rayleigh distribution with the second moment $\Omega = E[x^2] = 2\sigma^2$
Rice( $s, \sigma$ )	Rice distribution with noncentrality parameter $s$ and $\sigma$ . $s^2$ represent the specular component power
Rice( $\Omega, K$ ), Rice( $A, K$ )	Rice distribution with Rice factor $K = s^2/2\sigma^2$ and scale parameter $\Omega = A = s^2 + 2\sigma^2 = 2\sigma^2(K + 1)$ ( $\Omega$ is preferred over $A$ )

## 4 Machine learning, optimization theory, and statistical signal processing

### 4.1 Matrix Calculus (in denominator layout)

PS: Estimated terms come with a hat above of it

$\mathbf{g}, \nabla f, \frac{\partial f}{\partial \mathbf{w}}$	Gradient descent vector, used in the steepest descent method, also called gradient descent method or deterministic gradient method.
$\mathbf{g}$ if the gradient vector is $\nabla f$ (or $\hat{\mathbf{g}}$ if the gradient vector is $\mathbf{g}$ [25])	Stochastic gradient descent (SGD) vector, i.e., instantaneous approximation of gradient descent vector
$\mathbf{g}_x, \nabla_{\mathbf{w}} f, \frac{\partial f}{\partial \mathbf{w}}$	Gradient descent vector with respect $\mathbf{w}$ [9]
$\mathbf{J}, \frac{\partial \mathbf{y}^\top}{\partial \mathbf{x}}, \nabla \mathbf{y}^\top$ [25]	Jacobian matrix.
$\mathbf{H}, \frac{\partial^2 f}{\partial \mathbf{w}^2}, \nabla^2 f$ [25], $\nabla \nabla f$ [9]	Hessian matrix. The notation $\nabla^2$ is sometimes used in Matrix Calculus to denote the second-order vector derivative. However, it must be avoided since, in Vector Calculus, $\nabla^2$ also denotes the Laplacian operator which in turn may be scalar or vector Laplacian operator depending on whether $f$ is scalar- or vector-valued, respectively. Some discussion about can be found in [5–7]

### 4.2 Statistics: estimation and detection theory

$\mathbf{x}$	output
$\mathbf{w}$	Parameters
$p(\mathbf{x} \mid \mathbf{w}), l(\mathbf{x} \mid \mathbf{w})$ [31]	Likelihood function
$\ln p(\mathbf{x} \mid \mathbf{w})$	Log-likelihood function
$\Lambda(\mathbf{x})$ [31], $\frac{p(\mathbf{x} H_1)}{p(\mathbf{x} H_0)}$ [28, 31], $L(\mathbf{x})$ [13, 28]	Likelihood ratio function (also called likelihood ratio test (LRT) [28])
$\Lambda_l(\mathbf{x}), \mathcal{L}(\mathbf{x})$ [13], $l(\mathbf{x})$ [28]	Log-likelihood ratio (LLR [28]) function
$\hat{\rho}_{x,y}$	Estimated Pearson correlation coefficient between $x$ and $y$
$\mathcal{R}_k$	$k$ th Decision region
$x(t) \stackrel{m.s.e}{=} y(t)$	$x(t)$ equals $y(t)$ is the mean square error sense, that is $E \left[  x(t) - y(t) ^2 \right] = 0$
$x(t) = \text{l.i.m.} \sum_{i=1}^N x_i \phi_i(t)$ [43]	$\lim_{N \rightarrow \infty} E \left[ \left  x(t) - \sum_{i=1}^N x_i \phi_i(t) \right ^2 \right] = 0$ (l.i.m stands for “limit in the mean”). It is analogous to the $\stackrel{m.s.e}{=}$ notation, but denoting that they equal in the MSE sense only when $N \rightarrow \infty$

#### 4.3 Signals, (hyper)parameters, system performance, and criteria

$N$	Number of instances (or samples), i.e., $n \in \{1, 2, \dots, N\}$
$N_{\text{trn}}$	Number of instances in the training set, i.e., $n \in \{1, 2, \dots, N_{\text{trn}}\}$
$N_{\text{tst}}$	Number of instances in the test set, i.e., $n \in \{1, 2, \dots, N_{\text{tst}}\}$
$N_{\text{val}}$	Number of instances in the validation set, i.e., $n \in \{1, 2, \dots, N_{\text{val}}\}$
$N_e$	Number of epochs
$N_a$	Number of attributes
$K$ [25]	Number of classes (which is the number of outputs in multiclass problems). Use $k$ to iterate over it
$L$	Number of layers, i.e., the depth of the network. Use $l$ to iterate over it

$M_l, m_l$ [25], $J$ [25]	Number of neurons at the $l$ th layer. You might prefer $J$ in the case of the single-layer perceptron (use $j$ to iterate over it). If you want to iterate through it, a sensible variation of Haykin notation is $M_l$ , where $m_l$ can be used as an iterator. $m_0$ is the length of the input vector without the bias.
$\mathbf{x}(n), \mathbf{x}_n$	Input signal (in $\mathbb{R}^{N_a+1}$ )
$x_0(n)$	Dummy input of the bias, which is usually $\pm 1$ . $+1$ is preferred [9, 25].
$\varphi(\cdot)$ [25], $h(\cdot)$ [9]	Activation function
$\varphi'(v_{m_l}^{(l)}(n))$ [25], $\frac{\partial y_{m_l}^{(l)}(n)}{\partial v_{m_l}^{(l)}(n)}$ [25]	Partial derivative of the activation function with respect to $v_{m_l}^{(l)}(n)$ ( $m_l$ neuron at $l$ th layer)
$y_{m_l}^{(l)}(n), \varphi(v_{m_l}^{(l)}(n))$ [25], $t_{m_l}^{(l)}(n)$ [9]	Output signal (target) of the $m_l$ th neuron at the $l$ th layer
$\mathbf{y}^{(l)}(n)$	Output signal of the $l$ th layer
$\mathbf{y}(n), \mathbf{y}^{(L)}(n)$	Output of the neural network
$\mathbf{d}(n), \mathbf{d}_n$	Desired label (in case of supervised learning). For multiclass classification, one-hot encoding is usually used. For binary (scalar) classification, however antipodal encoding, i.e., $\{-1, 1\}$ is more recommended [25].
$e_{m_l}(n)$	Error signal of the neuron $m_l$ at the $l$ th layer
$\mathbf{e}(n), \mathbf{d}(n) - \mathbf{y}(n)$	Error signal
$\mathbf{w}_{m_l}^{(l)}(n), \boldsymbol{\theta}_{m_l}^{(l)}(n)$ $\begin{bmatrix} w_{m_l,0}^{(l)}(n) & w_{m_l,1}^{(l)}(n) & \dots & w_{m_l,m_{l-1}}^{(l)}(n) \end{bmatrix}$	Parameters, coefficients, or synaptic weights vector in the $l$ th layer. In the case of Single Layer Perceptrons or adaptive filters, the superscript is omitted
$w_{m_l,0}^{(l)}(n), b_{m_l}^{(l)}(n)$	Bias (the first term of the weight vector) of the $l$ th layer
$\mathbf{W}(n), [\mathbf{w}(1) \quad \mathbf{w}(2) \quad \dots \quad \mathbf{w}(N)]^\top$	Matrix of the synaptic weights
$\tilde{\mathbf{W}}(n)$	Matrix of the synaptic weights, but without the bias

$v_{m_l}^{(l)}(n), \mathbf{w}_{m_l}^{(l)\top}(n)\mathbf{y}_{m_{l-1}}^{(l-1)}(n)$	Induced local field or activation potential. At the first layer $\mathbf{y}_{m_0}^{(0)}(n) = \mathbf{x}(n)$ [9]
$\mathbf{v}^{(l)}(n), \mathbf{W}^{(l)}(n)\mathbf{y}_{m_{l-1}}^{(l-1)}(n)$	Vector of the local fields at the $l$ th layer
$\mathbf{w}^*, \mathbf{w}_o, \boldsymbol{\theta}^*, \boldsymbol{\theta}_o$	Optimum value of the parameters, coefficients, or synaptic weights vector ( $\mathbf{w}^*$ is also used [9] but it is not recommended as it may be confused with the conjugation operator)
$\delta_{m_l}^{(l)}(n), \frac{\partial \mathcal{E}(n)}{\partial v_{m_l}^{(l)}(n)}$	Local gradient of the $m_l$ th neuron of the $l$ th layer.
$\boldsymbol{\delta}^{(l)}(n)$	Vector of the local gradients of all neurons at the $l$ th layer
$\mathbf{X}, [\mathbf{x}(1) \ \mathbf{x}(2) \ \cdots \ \mathbf{x}(N)]$	Data matrix [25]
$\eta(n)$	Learning rate hyperparameter [25]
$\mathcal{R}$	Bayes risk or average risk [25]
$c_{ij}, C_{ij}$	Misclassification cost in deciding in favor of class $\mathcal{C}_i$ (represented in the subspace $\mathcal{H}_i$ ) when the $\mathcal{C}_j$ is the true class (used in Bayes classifiers/detectors) [13, 25]
$\mathcal{C}_k$ [25], $\mathcal{C}_k$ [9]	$k$ th class
$\mathcal{T}$ [25], $\mathbb{X}$ [23]	Training set, i.e., the set $\{\mathbf{x}(n), d(n)\}$ that is used in the training phase.
$\mathcal{H}_k$	Subspace of the training vector belonging to the class $\mathcal{C}_k$
$\mathcal{H}$	Complete space of the input vector, i.e., $\mathcal{H}_1 \cup \mathcal{H}_2 \cup \cdots \mathcal{H}_K$
$\mathcal{X}$ [25]	Set of all vectors in the training, batch, validation, or test dataset that were misclassified
$\mathcal{E}(\mathbf{w}), \mathcal{E}(\mathbf{w}(n)), \mathcal{E}(n)$	Cost function or objective function (the way it is written depends on the purpose of the text)
$J(\mathbf{w}), J(\mathbf{w}(n)), J(n)$	Alternative to the cost function
$\Delta \mathcal{E}(\mathbf{w}(n)), \Delta \mathcal{E}(n), \mathcal{E}(\mathbf{w}(n+1)) - \mathcal{E}(\mathbf{w}(n))$	Cost function or objective function (the way it is written depends on the purpose of the text)
$\mathcal{E}_{\text{av}}(\cdot)$ [25]	Error energy averaged over the training sample or the empirical risk

$\rho$	Distance of the margin of separation between two classes (Support Vector Machine, SVM)
$g(\cdot)$	Discriminant function, i.e., $g(\mathbf{w}^*) = 0$

## 5 Linear Algebra

### 5.1 Common matrices and vectors

<b>W, D</b>	Diagonal matrix
<b>P</b>	Projection matrix; Permutation matrix
<b>S</b>	Symmetric matrix
<b>J</b>	Jordan matrix
<b>L</b>	Lower matrix
<b>U</b>	Upper matrix; Unitary matrix
<b>C</b>	Cofactor matrix
<b>C<sub>A</sub>, cof (A)</b>	Cofactor matrix of <b>A</b>
<b>S</b>	Symmetric matrix
<b>Q</b>	Orthogonal matrix
<b>I<sub>N</sub></b>	$N \times N$ -dimensional identity matrix
<b>0<sub>M×N</sub></b>	$M \times N$ -dimensional null matrix
<b>0<sub>N</sub></b>	$N$ -dimensional null vector
<b>1<sub>M×N</sub></b>	$M \times N$ -dimensional ones matrix
<b>1<sub>N</sub></b>	$N$ -dimensional ones vector
<b>0</b>	Null matrix, vector, or tensor (dimensionality understood by context)
<b>1</b>	Ones matrix, vector, or tensor (dimensionality understood by context)

### 5.2 Indexing

$x_{i_1, i_2, \dots, i_N}, [\mathcal{X}]_{i_1, i_2, \dots, i_N}$	Element in the position $(i_1, i_2, \dots, i_N)$ of the tensor $\mathcal{X}$
$\mathcal{X}^{(n)}$	$n$ th tensor of a nontemporal sequence
$\mathbf{x}_n, \mathbf{X}_{:n}$	$n$ th column of the matrix $X$
$\mathbf{x}_{n:}$	$n$ th row of the matrix $X$
$\mathbf{x}_{i_1, \dots, i_{n-1}, :, i_{n+1}, \dots, i_N}$	Mode- $n$ fiber of the tensor $\mathcal{X}$
$\mathbf{x}_{:, i_2, i_3}$	Column fiber (mode-1 fiber) of the thrid-order tensor $\mathcal{X}$

$\mathbf{x}_{i_1, :, i_3}$	Row fiber (mode-2 fiber) of the thrid-order tensor $\mathcal{X}$
$\mathbf{x}_{i_1, i_2, :}$	Tube fiber (mode-3 fiber) of the thrid-order tensor $\mathcal{X}$
$\mathbf{X}_{i_1, :, :}$	Horizontal slice of the thrid-order tensor $\mathcal{X}$
$\mathbf{X}_{:, i_2, :}$	Lateral slices slice of the thrid-order tensor $\mathcal{X}$
$\mathbf{X}_{i_3}, \mathbf{X}_{:, :, i_3}$	Frontal slices slice of the thrid-order tensor $\mathcal{X}$

### 5.3 General operations

$\langle \mathbf{a}, \mathbf{b} \rangle, \mathbf{a}^\top \mathbf{b}, \mathbf{a} \cdot \mathbf{b}$	Inner or dot product
$\mathbf{a} \circ \mathbf{b}, \mathbf{a} \mathbf{b}^\top$	Outer product
$\otimes$	Kronecker product
$\odot$	Hadamard (or Schur) (elementwise) product
$\cdot^{\odot n}$	$n$ th-order Hadamard power
$\cdot^{\odot \frac{1}{n}}$	$n$ th-order Hadamard root
$\oslash$	Hadamard (or Schur) (elementwise) division
$\diamond$	Khatri-Rao product
$\otimes$	Kronecker Product
$\times_n$	$n$ -mode product

### 5.4 Operations with matrices and tensors

$\mathbf{A}^{-1}$	Inverse matrix
$\mathbf{A}^+, \mathbf{A}^\dagger$	Moore-Penrose left pseudoinverse
$\mathbf{A}^\top, \mathbf{A}^T, \mathbf{A}^t, \mathbf{A}'$ [38]	Transpose
$\mathbf{A}^{-\top}$	Transpose of the inverse, i.e., $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$ [22, 34]
$\mathbf{A}^*$	Complex conjugate
$\mathbf{A}^H$	Hermitian
$\ \mathbf{A}\ _F$	Frobenius norm
$\ \mathbf{A}\ $	Matrix norm
$ \mathbf{A} , \det(\mathbf{A})$	Determinant
$\text{diag}(\mathbf{A})$	The elements in the diagonal of $\mathbf{A}$
$\text{vec}[\mathbf{A}]$	Vectorization: stacks the columns of the matrix $\mathbf{A}$ into a long column vector

$\text{vec}_d [\mathbf{A}]$	Extracts the diagonal elements of a square matrix and returns them in a column vector
$\text{vec}_l [\mathbf{A}]$	Extracts the elements strictly below the main diagonal of a square matrix in a column-wise manner and returns them into a column vector
$\text{vec}_u [\mathbf{A}]$	Extracts the elements strictly above the main diagonal of a square matrix in a column-wise manner and returns them into a column vector
$\text{vec}_b [\mathbf{A}]$	Block vectorization operator: stacks square block matrices of the input into a long block column matrix
$\text{unvec} (\mathbf{A})$	Reshapes a column vector into a matrix
$\text{tr}\{\mathbf{A}\}$	trace
$\mathbf{X}_{(n)}$	$n$ -mode matricization of the tensor $\mathcal{X}$

## 5.5 Operations with vectors

$\ \mathbf{a}\ $	$l_1$ norm, 1-norm, or Manhattan norm
$\ \mathbf{a}\ , \ \mathbf{a}\ _2$	$l_2$ norm, 2-norm, or Euclidean norm
$\ \mathbf{a}\ _p$	$l_p$ norm, $p$ -norm, or Minkowski norm
$\ \mathbf{a}\ _\infty$	$l_\infty$ norm, $\infty$ -norm, or Chebyshev norm
$\text{diag} (\mathbf{a})$	Diagonalization: a square, diagonal matrix with entries given by the vector $\mathbf{a}$

## 5.6 Decompositions

$\mathbf{\Lambda}$	Eigenvalue matrix [40]
$\mathbf{Q}$	Eigenvectors matrix; Orthogonal matrix of the QR decomposition[40]
$\mathbf{R}$	Upper triangular matrix of the QR decomposition[40]
$\mathbf{U}$	Left singular vectors[40]
$\mathbf{U}_r$	Left singular nondegenerated vectors
$\mathbf{\Sigma}$	Singular value matrix
$\mathbf{\Sigma}_r$	Singular value matrix with nonzero singular values in the main diagonal

$\Sigma^+$	Singular value matrix of the pseudoinverse [40]
$\Sigma_r^+$	Singular value matrix of the pseudoinverse with nonzero singular values in the main diagonal
$\mathbf{V}$	Right singular vectors [40]
$\mathbf{V}_r$	Right singular nondegenerated vectors
$\text{eig}(\mathbf{A})$	Set of the eigenvalues of $\mathbf{A}$ [14, 31, 34]
$[[\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots]]$	CANDECOMP/PARAFAC (CP) decomposition of the tensor $\mathcal{X}$ from the outer product of column vectors of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$
$[[\lambda; \mathbf{A}, \mathbf{B}, \mathbf{C}, \dots]]$	Normalized CANDECOMP/PARAFAC (CP) decomposition of the tensor $\mathcal{X}$ from the outer product of column vectors of $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$

## 5.7 Spaces and sets

### 5.7.1 Common spaces and sets

$\mathbb{R}$	Set of real numbers
$[a, b]$	Closed interval of a real set from $a$ to $b$
$(a, b)$	Opened interval of a real set from $a$ to $b$
$[a, b), (a, b]$	Half-opened intervals of a real set from $a$ to $b$
$\mathbb{C}$	Set of complex numbers
$\mathbb{I}, j\mathbb{R}$	Set of imaginary numbers
$\mathbb{Q}$	Set of rational number
$\mathbb{R} \setminus \mathbb{Q}$	Set of irrational number
$\mathbb{Z}$	Set of integer number
$\mathbb{N}$	Set of natural numbers
$\{1, 2, \dots, n\}$	Discrete set containing the integer elements $1, 2, \dots, n$
$\mathbb{B} = \{0, 1\}$	Boolean set
$\emptyset$	Empty set
$\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$	Real or complex space (field)



$\mathbb{K}^{I_1 \times I_2 \times \dots \times I_N}$	$I_1 \times I_2 \times \dots \times I_N$ -dimensional real (or complex) space
$\mathbb{K}_+$	Nonnegative real (or complex) space [10]
$\mathbb{K}_{++}$	Positive real (or complex) space, i.e., $\mathbb{K}_{++} = \mathbb{K}_+ \setminus \{\mathbf{0}\}$ [10]
$U$	Universe
$2^A$	Power set of $A$

### 5.7.2 Convex sets (or spaces)

$\mathbb{S}^n$ [16], $\mathcal{S}^n$ [10]	Conic set of the symmetric matrices in $\mathbb{R}^{n \times n}$
$\mathbb{S}^{n\perp}$ [16]	Conic set of the skew-symmetric (also called antisymmetric) matrices in $\mathbb{R}^{n \times n}$
$\mathbb{S}_+^n, \mathcal{S}_+^n$	Conic set of the symmetric positive semidefinite matrices in $\mathbb{R}^{n \times n}$ [10]
$\mathbb{S}_{++}^n, \mathcal{S}_{++}^n$	Conic set of the symmetric positive definite matrices in $\mathbb{R}^{n \times n}$ , i.e., $\mathbb{S}_{++}^n = \mathbb{S}_+^n \setminus \{\mathbf{0}\}$ [10]
$\mathbb{H}^n$	Set of all hermitian matrices in $\mathbb{C}^{n \times n}$
$\text{conv } C$	Convex hull
$\text{aff } C$	Affine hull
$\mathcal{R}$	Ray
$\mathcal{H}$	Hyperplane
$\mathcal{H}_+, \mathcal{H}_-$	Positive/negative halfspace
$B(\mathbf{x}_c, r)$	Euclidean ball with radius $r$ and centered at $\mathbf{x}_c$
$\mathcal{E}$	Ellipsoid
$C$	Norm cone
$K$	Proper cone
$K^*$	Dual cone
$\mathcal{P}$	Polyhedra
$S$	Simplex
$C_\alpha$	$\alpha$ -sublevel set
$\text{epi } f$	Epigraph of the function $f$
$\text{hypo } f$	Hypograph of the function $f$

### 5.7.3 Spaces from matrices or vectors

$\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$	Vector space spanned by the argument vectors [22]
$C(\mathbf{A}), \text{columnspace}(\mathbf{A}), \text{range}(\mathbf{A}), \text{span}\{\mathbf{A}\}, \text{image}(\mathbf{A})$	Columnspace, range or image, i.e., the space $\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ , where $\mathbf{a}_i$ is the $i$ th column vector of the matrix $\mathbf{A}$ [32, 40]
$C(\mathbf{A}^H)$	Row space (also called left column space) [32, 40]
$N(\mathbf{A}), \text{nullspace}(\mathbf{A}), \text{null}(\mathbf{A}), \text{kernel}(\mathbf{A})$	Nullspace (or kernel space) [32, 40, 41]
$N(\mathbf{A}^H)$	Left nullspace
$\text{rank } \mathbf{A}$	Rank, that is, $\dim(\text{span}\{\mathbf{A}\}) = \dim(C(\mathbf{A}))$ [32]
$\text{nullity}(\mathbf{A})$	Nullity of $\mathbf{A}$ , i.e., $\dim(N(\mathbf{A}))$

## 5.8 Set operations

$A + B$	Set addition (Minkowski sum), i.e., $\{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} = \mathbf{x} + \mathbf{y}, \forall \mathbf{x} \in \mathcal{X} \wedge \mathbf{y} \in \mathcal{Y}\}$ [29]
$A - B$	Minkowski difference, i.e., $\{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} = \mathbf{x} - \mathbf{y}, \forall \mathbf{x} \in \mathcal{X} \wedge \mathbf{y} \in \mathcal{Y}\}$
$A \ominus B$	Pontryagin difference, i.e., $\{\mathbf{v} \in \mathbb{R}^n \mid \mathbf{v} + \mathbf{y} \in \mathcal{X}, \forall \mathbf{y} \in \mathcal{Y}\}$ [29]
$A \setminus B, A - B$	Set difference or set subtraction, i.e., $A \setminus B = \{x \mid x \in A \wedge x \notin B\}$ the set containing the elements of $A$ that are not in $B$ [37]
$A \cup B$	Set of union
$A \cap B$	Set of intersection
$A \times B$	Cartesian product
$A^n$	$\underbrace{A \times A \times \dots \times A}_{n \text{ times}}$
$A^\perp$	Orthogonal complement of $A$ , e.g., $N(\mathbf{A}) = C(\mathbf{A}^\top)^\perp$ [10]
$\mathbf{a} \perp \mathbf{b}$	$\mathbf{a}$ is orthogonal to $\mathbf{b}$
$\mathbf{a} \not\perp \mathbf{b}$	$\mathbf{a}$ is not orthogonal to $\mathbf{b}$
$A \oplus B$	Direct sum, i.e., each $\mathbf{v} \in \{\sum \mathbf{a}_i \mid \mathbf{a}_i \in S_i, i = 1, \dots, k\}$ has a unique representation of $\sum \mathbf{a}_i$ with $\mathbf{a}_i \in S_i$ . That is, they expand to a space. Note that $\{S_i\}$ might not be orthogonal each other [22]

$A \overset{\perp}{\oplus} B$	Direct sum of two spaces that are orthogonal and span a $n$ -dimensional space, e.g., $C(\mathbf{A}^\top) \overset{\perp}{\oplus} C(\mathbf{A}^\top)^\perp = \mathbb{R}^n$ (this decomposition of $\mathbb{R}^n$ is called the orthogonal decomposition induced by $\mathbf{A}$ ) [10]
$\bar{A}, A^c$	Complement set (given $U$ )
$\#A,  A $	Cardinality of $A$
$a \in A$	$a$ is element of $A$
$a \notin A$	$a$ is not element of $A$

## 5.9 Inequalities

$\mathcal{X} \preceq 0$	Nonnegative tensor
$\mathbf{a} \preceq_K \mathbf{b}$	Generalized inequality meaning that $\mathbf{b} - \mathbf{a}$ belongs to the conic subset $K$ in the space $\mathbb{R}^n$ [10]
$\mathbf{a} \prec_K \mathbf{b}$	Strict generalized inequality meaning that $\mathbf{b} - \mathbf{a}$ belongs to the interior of the conic subset $K$ in the space $\mathbb{R}^n$ [10]
$\mathbf{a} \preceq \mathbf{b}$	Generalized inequality meaning that $\mathbf{b} - \mathbf{a}$ belongs to the nonnegative orthant conic subset, $\mathbb{R}_+^n$ , in the space $\mathbb{R}^n$ . [10]
$\mathbf{a} \prec \mathbf{b}$	Strict generalized inequality meaning that $\mathbf{b} - \mathbf{a}$ belongs to the positive orthant conic subset, $\mathbb{R}_{++}^n$ , in the space $\mathbb{R}^n$ [10]
$\mathbf{A} \preceq_K \mathbf{B}$	Generalized inequality meaning that $\mathbf{B} - \mathbf{A}$ belongs to the conic subset $K$ in the space $\mathbb{S}^n$ [10]
$\mathbf{A} \prec_K \mathbf{B}$	Strict generalized inequality meaning that $\mathbf{B} - \mathbf{A}$ belongs to the interior of the conic subset $K$ in the space $\mathbb{S}^n$ [10]
$\mathbf{A} \preceq \mathbf{B}$	Generalized inequality meaning that $\mathbf{B} - \mathbf{A}$ belongs to the positive semidefinite conic subset, $\mathbb{S}_+^n$ , in the space $\mathbb{S}^n$ [10]
$\mathbf{A} \prec \mathbf{B}$	Strict generalized inequality meaning that $\mathbf{B} - \mathbf{A}$ belongs to the positive orthant conic subset, $\mathbb{S}_{++}^n$ , in the space $\mathbb{S}^n$ [10]

## 6 Communication systems

### 6.1 Common symbols

$B$	One-sided bandwidth of the base-band signal, in Hz
$W$	One-sided bandwidth of the base-band signal, in rad/s
$N_0$	Noise density, in ???
$x_i$	Real or in-phase part of $x$
$x_q$	Imaginary or quadrature part of $x$
$f_c, f_{RF}$	Carrier frequency (in Hertz)
$f_L$	Carrier frequency in L-band (in Hertz)
$f_{IF}$	Intermediate frequency (in Hertz)
$f_s$	Sampling frequency or sampling rate (in Hertz)
$T_s$	Sampling time interval/duration/period
$R$	Bit rate
$T$	Bit interval/duration/period
$T_c$	Chip interval/duration/period
$T_{sy}, T_{sym}$	Symbol/signaling[35] interval/duration/period
$s_{RF}$	Transmitted signal in RF
$s_{FI}$	Transmitted signal in FI
$s, s_l$	Lowpass (or baseband) equivalent signal or envelope complex of transmitted signal
$r_{RF}$	Received signal in RF
$r_{FI}$	Received signal in FI
$r, r_l$	Lowpass (or baseband) equivalent signal or envelope complex of received signal
$\phi$	Signal phase
$\phi_0$	Initial phase
$\eta_{RF}, w_{RF}$	Noise in RF
$\eta_{FI}, w_{FI}$	Noise in FI
$\eta, w$	Noise in baseband
$\tau$	Timing delay
$\Delta\tau$	Timing error (delay - estimated)
$\varphi$	Phase offset
$\Delta\varphi$	Phase error (offset - estimated)
$f_d$	Linear Doppler frequency

$\Delta f_d$	Frequency error (Doppler frequency - estimated)
$\nu$	Angular Doppler frequency
$\Delta \nu$	Frequency error (Doppler frequency - estimated)
$\gamma, A$	Transmitted signal amplitude
$\gamma_0, A_0$	Combined effect of the path loss and antenna gain

## 6.2 Fading multipath channels

$t \xleftrightarrow{\mathcal{F}} \lambda$ [35]	Support temporal of the signal. $\lambda$ is obtained after taking the Fourier transform on $t$ .
$\tau \xleftrightarrow{\mathcal{F}} f$ [35]	Second support temporal of the signal ( $c(t)$ varies with the input at the time $\tau$ ). $f$ is obtained after taking the Fourier transform on $\tau$ .
$c(t, \tau)$ [35]	Complex envelope of the channel response at the time $t$ due to an impulse applied at the $t - \tau$
$C(f, t)$ [35]	Transfer function of $c(t, \tau)$ in $\tau$
$\alpha(t, \tau)$ [35]	Attenuation of $c(t, \tau)$ , i.e., $c(t, \tau) = \alpha(t, \tau)e^{e\pi f_c \tau}$
$R_c(\tau_1, \tau_2, \Delta t)$ [35]	Autocorrelation function of $c(t, \tau)$ , i.e., $R_c(\tau_1, \tau_2, \Delta t) = \text{E} [c^*(t, \tau_1), c^*(t + \Delta t, \tau_2)]$
$R_c(\tau, \Delta t)$ [35]	Autocorrelation function of $c(t, \tau)$ assuming uncorrelated scattering
$R_c(\tau), R_c(\tau, \Delta t) _{\Delta t=0}$ [35]	Multipath intensity profile or delay power spectrum
$R_C(\Delta f, \Delta t), R_C(f_1, f_2; \Delta t), \text{E} [C(f_1, t), C(f_2, t + \Delta t)], \mathcal{F}_\tau \{R_c(\tau, \Delta t)\}$ [21]	Spaced-frequency, spaced-time correlation function ( $\Delta f = f_2 - f_1$ )
$R_C(\Delta f), R_C(\Delta f, \Delta t) _{\Delta t=0}$ [35], $\mathcal{F} \{R_c(\tau)\}$ [21]	Spaced-frequency correlation function
$(\Delta f)_c$	Coherence bandwidth of $c(t)$ , that is, the frequency interval in which $R_C(\Delta f)$ is nonzero [35]
$T_m$	Multipath spread of the channel, that is, the time interval in which $R_c(\tau)$ is nonzero ( $T_m \approx 1/(\Delta f)_c$ ) [35]

$R_C(\Delta t), R_C(\Delta f, \Delta t) _{\Delta f=0}$	Spaced-time correlation function [35]
$S_C(\lambda)$ [35], $\mathcal{F}\{R_C(\Delta t)\}$ [21]	Doppler power spectrum
$(\Delta t)_c$	Coherence time of $c(t)$ , that is, the time interval in which $R_C(\Delta t)$ is nonzero [35]
$B_m$	Multipath spread of the channel, that is, the frequency interval in which $S_c(\lambda)$ is nonzero ( $B_d \approx 1/(\Delta t)_c$ ) [35]
$S_C(\tau, \lambda)$ [35], $\mathcal{F}_{\Delta f, \Delta t}\{R_C(\Delta f, \Delta t)\}$ [21]	Scattering function

## 7 Discrete mathematics

### 7.1 Quantifiers, inferences

$\forall$	For all (universal quantifier) [24]
$\exists$	There exists (existential quantifier) [24]
$\nexists$	There does not exist [24]
$\exists!$	There exists an unique [24]
$\exists_n$	There exists exactly $n$ [37]
$\in$	Belongs to [24]
$\notin$	Does not belong to [24]
$\because$	Because [24]
$ , :$	Such that, sometimes that parentheses is used [24]
$,, (\cdot)$	Used to separate the quantifier with restricted domain from its scope, e.g., $\forall x < 0 (x^2 > 0)$ or $\forall x < 0, x^2 > 0$ [24]
$\therefore$	Therefore [24]

### 7.2 Propositional Logic

$\neg a$	Logical negation of $a$ [37]
$a \wedge b$	Conjunction (logical AND) operator between $a$ and $b$ [37]
$a \vee b$	Disjunction (logical OR) operator between $a$ and $b$ [37]
$a \oplus b$	Exclusive OR (logical XOR) operator between $a$ and $b$ [37]

$a \rightarrow b$	Implication (or conditional) statement[37]
$a \leftrightarrow b$	Bi-implication (or biconditional) statement, i.e., $(a \rightarrow b) \wedge (b \rightarrow a)$ [37]
$a \equiv b, a \iff b, a \Leftrightarrow b$	Logical equivalence, i.e., $a \leftrightarrow b$ is a tautology[37]

### 7.3 Operations

$ a $	Absolute value of $a$
$\log$	Base-10 logarithm or decimal logarithm
$\ln$	Natural logarithm
$\operatorname{Re}\{x\}$	Real part of $x$
$\operatorname{Im}\{x\}$	Imaginary part of $x$
$\angle \cdot$	Phase (complex argument)
$x \bmod y$	Remainder, i.e., $x - y\lfloor x/y \rfloor$ , for $y \neq 0$
$x \operatorname{div} y$	Quotient [37]
$x \equiv y \pmod{m}$	Congruent, i.e., $m \mid (x - y)$ [37]
$\operatorname{frac}(x)$	Fractional part, i.e., $x \bmod 1$ [24]
$a \setminus b$ [24, Section 4.1], $a \mid b$ [37]	$b$ is a positive integer multiple of $a \in \mathbb{Z}$ , i.e., $\exists! n \in \mathbb{Z}_{++} \mid b = na$
$a \nmid b$ [24, Section 4.1], $a \nmid b$ [37]	$b$ is not a positive integer multiple of $a \in \mathbb{Z}$ , i.e., $\nexists n \in \mathbb{Z}_{++} \mid b = na$
$\lceil \cdot \rceil$	Ceiling operation [24]
$\lfloor \cdot \rfloor$	Floor operation [24]

## 8 Vector Calculus

$\nabla f$ [39], $\operatorname{grad} f$ [36]	Vector differential operator (Nabla symbol), i.e., $\nabla f$ is the gradient of the scalar-valued function $f$ , i.e., $f : \mathbb{R}^n \rightarrow \mathbb{R}$
$t, (u, v)$	Parametric variables commonly used, $t$ for one variable, $(u, v)$ for two variables[39]
$\mathbf{l}(x, y, z)$ [36], $\mathbf{r}(x, y, z)$ [39], $x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$	Vector position, i.e., $(x, y, z)$ .
$\mathbf{l}(t)$	Vector position parametrized by $t$ , i.e., $(x(t), y(t), z(t))$ [36, 39]

$\mathbf{l}'(t), d\mathbf{l}/dt$	First derivative of $\mathbf{l}(t)$ , i.e., the tangent vector of the curve $(x(t), y(t), z(t))$ [39]
$\mathbf{u}(t)$ [30] $\mathbf{T}(t)$ [39], $d\mathbf{l}(t)$ [36]	Tangent unit vector of $\mathbf{l}(t)$ , i.e., $\mathbf{u}(t) = \mathbf{l}'(t)/ \mathbf{l}'(t) $
$\mathbf{n}(t), \left( \frac{y'(t)}{ \mathbf{l}'(t) }, -\frac{x'(t)}{ \mathbf{l}'(t) } \right)$	Normal vector of $\mathbf{l}(t)$ , i.e., $\mathbf{n}(t) \perp \mathbf{T}(t)$ [39]
$C$	Contour that traveled by $\mathbf{l}(t)$ , for $a \leq t \leq b$ [39]
$L, L(C)$	Total length of the contour $C$ (which can be defined the vector $\mathbf{l}$ , parametrized by $t$ ), i.e., $L_C = \int_a^b  \mathbf{l}'(t)  dt$ [39]
$s(t)$	Length of the arc, which can be defined by the vector $\mathbf{l}$ and $t$ , that is, $s(t) = \int_a^t  \mathbf{l}'(u)  du$ ( $s(b) = L$ ) [39]
$ds$	Differential operator of the length of the contour $C$ , i.e., $ds =  \mathbf{l}'(t)  dt$ [39]
$\int_C f(\mathbf{l}) ds, \int_a^b f(\mathbf{l}(t)) \mathbf{l}'(t)  dt$	Line integral of the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ along the contour $C$ . In the context of integrals in the complex plane, it is also called “contour integral”
$\theta$ [36]	Angle between the contour $C$ and the vector field $\mathbf{F}$
$\int_C \mathbf{F} \cdot d\mathbf{l}, \int_a^b \mathbf{F}(\mathbf{l}(t)) \cdot \mathbf{l}'(t) dt$ [8, 39], $\int_C \mathbf{F} \cdot \mathbf{u} ds, \int_C \mathbf{F} \cos(\theta) ds$ [36]	Line integral of vector field $\mathbf{F}$ along the contour $C$
$\int_C \mathbf{F} \cdot d\mathbf{u}$ [36]	In the field of electromagnetics, it is common to apply the line integral between the vector field $\mathbf{F}$ and the unit vector $\mathbf{u}(t)$ . Therefore, this line integral may appear as well
$\int_a^b \mathbf{F}, \int_a^b \mathbf{F} \cdot d\mathbf{l}$	Alternative notation to the line integral, where the parametric variable $t$ goes from $a$ to $b$ , making $r$ goes from $\mathbf{l}(a) = \mathbf{a}$ to $\mathbf{l}(b) = \mathbf{b}$ [8]
$\oint_C, \oint_C$	Line integral along the closed contour $C$ . The arrow indicates the contour integral orientation, which is counter-clockwise, by default. In the context of integrals in the complex plane, it is also called “closed contour integral”.



$\oint_S$	Surface integral over the closed surface $S$
$\mathbf{l}(u, v)$	Vector position $(x(u, v), y(u, v), z(u, v))$ parametrized by $(u, v)$
$\mathbf{l}_u$	$(\partial x / \partial u, \partial y / \partial u, \partial z / \partial u)$
$\mathbf{l}_v$	$(\partial x / \partial v, \partial y / \partial v, \partial z / \partial v)$
$dA$	Differential operator of a 2D area (denoted by $D$ or $R$ ) in the $\mathbb{R}^2$ domain. This differential operator can be solved in different ways (rectangular, polar, cylindric, etc) [39]
$D, R$	Integration domain in which $dA$ is integrated, i.e., $\iint_D f dA$ . $R$ is preferred when the integration domain is a rectangle, while $D$ is used when it has nonrectangular shape [39]
$S$	Smooth surface $S \subset \mathbb{R}^3$ , i.e., a 2D area in a 3D space
$dS,  \mathbf{l}_u \times \mathbf{l}_v  dA$	Differential operator of a 2D area in a 3D domain (an surface). Note that $dS =  \mathbf{l}_u \times \mathbf{l}_v  dA$ should be accompanied with the change of the integration interval (from $S$ to $D$ )
$A(S), \iint_S dS, \iint_D  \mathbf{l}_u \times \mathbf{l}_v  dA$	Area of the surface $S$ parametrized by $(u, v)$ , in which $dA$ is the area defined in the $D$ domain (which is form by the $u$ -by- $v$ graph)
$dV$	Differential operator of a shape volume (denoted by $E$ ) in $\mathbb{R}^3$ domain, i.e., $\iiint_E dV = V$
$E$	Integration domain in which $dV$ is integrated, i.e., $\iiint_E f dV$ [39]
$V, \iint_D f dA, \iiint_E f dV$	Volume of the function $f$ over the regions $D$ (in the case of double integrals) or $E$ (in the case of triple integrals)
$\iint_S f dS, \iint_D f  \mathbf{l}_u \times \mathbf{l}_v  dA$	Surface integral over $S$
$\mathbf{n}(u, v), \frac{\mathbf{l}_u(u, v) \times \mathbf{l}_v(u, v)}{ \mathbf{l}_u(u, v) \times \mathbf{l}_v(u, v) }$	Normal vector of of the smooth surface $S$
$\iint_S \mathbf{F} \cdot \mathbf{n} dS, \iint_S \mathbf{F} \cdot d\mathbf{S}, \iint_D \mathbf{F} \cdot (\mathbf{l}_u \times \mathbf{l}_v) dA$	Flux integral of vector field $\mathbf{F}$ through the smooth surface $S$ ( $\mathbf{n} dS \triangleq d\mathbf{S}$ )

$\oint_S \mathbf{F} \cdot \mathbf{n} dS, \oint_S \mathbf{F} \cdot d\mathbf{S},$ $\iint_D \mathbf{F} \cdot (\mathbf{l}_u \times \mathbf{l}_v) dA$	Flux integral of vector field $\mathbf{F}$ through the smooth and closed surface $S$ ( $\mathbf{n} dS \triangleq d\mathbf{S}$ )
$\nabla \times \mathbf{F}, \text{curl } \mathbf{F}$	Curl (rotacional) of the vector field $\mathbf{F}$
$\nabla \cdot \mathbf{F}, \text{div } \mathbf{F}$	Divergence of the vector field $\mathbf{F}$
$\nabla^2 f, \nabla \cdot (\nabla f), \Delta f,$ $\partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 + \partial^2 f / \partial z^2$	Scalar Laplacian operator (performed on a scalar-valued function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ )
$\nabla^2 \mathbf{F}, \nabla \times \nabla \times \mathbf{F} - \nabla(\nabla \cdot \mathbf{F}), \Delta \mathbf{F},$ $(\partial^2 \mathbf{F} / \partial x^2, \partial^2 \mathbf{F} / \partial y^2, \partial^2 \mathbf{F} / \partial z^2)$	Vector Laplacian operator (performed on a vector field, i.e., a vector-valued function, $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ). $\nabla^2$ denotes the scalar (vector) Laplacian if the function is scalar-valued (vector-valued). The notation $\Delta$ must be avoided as it is overused in many contexts

## 9 Electromagnetic waves

$\Phi$	Electric flux (scalar) (in V m)
$\mathbf{H}$	Magnetic field vector (in A/m)
$\mathbf{B}$	Magnetic flux density vector (in Wb/m <sup>2</sup> = T)
$\Phi$ <a href="#">[15]</a>	Magnetic flux
$q_f, q_{\text{free}}, Q_{\text{free}}$ <a href="#">[19]</a>	Free electric charge (in C)
$q_b, q_{\text{bound}}, Q_{\text{bound}}$ <a href="#">[19]</a>	Bound electric charge (in C)
$q, q_f + q_b$	Electric charge (in C)
$\rho_f$ <a href="#">[1]</a> , $\rho_{\text{free}}$ <a href="#">[19]</a>	Free electric charge density
$\rho_b$ <a href="#">[1]</a> , $\rho_{\text{bound}}$ <a href="#">[19]</a>	Electric charge density
$\rho, \rho_f + \rho_b$	Electric charge density (it can be in C/m <sup>3</sup> , C/m <sup>2</sup> or C/m depending whether it is a volume, surface, or line shapes)
$\mathbf{f}$ <a href="#">[36]</a> , $\mathbf{F}$ <a href="#">[2]</a>	Electrostatic force (Coulomb force), (in kg m/s <sup>2</sup> ).
$\varepsilon$	Electric permittivity (in F/m). If the medium is isotropic, it is a scalar. If it is anisotropic, it is a tensor. <a href="#">[36]</a>
$\varepsilon_r$	Relative electric permittivity or dielectric constant (in F/m) <a href="#">[36]</a>
$\varepsilon_0$	Electric permittivity in vacuum, $8.854 \times 10^{-12}$ F/m <a href="#">[36]</a>

$\mathbf{E}$	Electric field vector (in V/m)
$\sigma$	Electric conductivity (in S/m)
$\mathbf{J}$	Electric current density vector (in A/m <sup>2</sup> )
$\mathbf{J}_m$ [15]	Magnetization current density vector (in A/m <sup>2</sup> )
$\mathbf{D}$	Electric flux density, electric displacement, or electric induction vector (in C/m <sup>2</sup> )
$U$	Electric potential energy
$V$ [3, 15], $\Phi$ [36]	Electric potential (in voltage, V). However, keep in mind that there is a subtle difference between both definitions [4]
$\Phi_E$ [20], $\oint_S \mathbf{E} d\mathbf{S}$	Electric flux (in V m)
$\Phi_D$ [19], $\Psi$ [36], $\oint_S \mathbf{D} d\mathbf{S}$	Electric flux ( $\mathbf{D}$ -field flux)
$\mathbf{P}$	Electric polarization of the material (in C/m <sup>2</sup> )
$\chi_e$	Electric susceptibility (for linear and isotropic materials)
$\mu$	Magnetic permeability
$\mu_0$	Magnetic permeability in vacuum

## 10 Generic mathematical symbols

$\blacksquare$	Q.E.D.
$\triangleq$	Equal by definition
$:=, \leftarrow$	Assignment [37]
$\neq$	Not equal
$\infty$	Infinity
$j$	$\sqrt{-1}$

## 11 Abbreviations

wrt.	With respect to
st.	Subject to
iff.	If and only if
EVD	Eigenvalue decomposition, or eigen-decomposition [32]
SVD	Singular value decomposition
CP	CANDECOMP/PARAFAC

## References

- [1] URL: [https://en.wikipedia.org/wiki/Electric\\_displacement\\_field#Definition](https://en.wikipedia.org/wiki/Electric_displacement_field#Definition).
- [2] URL: [https://en.wikipedia.org/wiki/Coulomb%27s\\_law](https://en.wikipedia.org/wiki/Coulomb%27s_law).
- [3] URL: [https://en.wikipedia.org/wiki/Electric\\_potential](https://en.wikipedia.org/wiki/Electric_potential).
- [4] URL: <https://physics.stackexchange.com/a/300937/368410>.
- [5] Libavius (<https://math.stackexchange.com/users/1020990/libavius>). *Which is the correct vector calculus notation for the Hessian?* Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/4560326> (version: 2023-02-15). eprint: <https://math.stackexchange.com/q/4560326>. URL: <https://math.stackexchange.com/q/4560326>.
- [6] maple (<https://math.stackexchange.com/users/51601/maple>). *Does the symbol  $\nabla^2$  has the same meaning in Laplace Equation and Hessian Matrix?* Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/1353761> (version: 2022-07-29). eprint: <https://math.stackexchange.com/q/1353761>. URL: <https://math.stackexchange.com/q/1353761>.
- [7] Rubem Pacelli (<https://math.stackexchange.com/users/817590/rubem-pacelli>). *Ambiguity over the notation  $\nabla^2$ : vector Laplacian operator (Vector Calculus) vs. second directional derivative (Matrix Calculus)*. Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/4693212> (version: 2023-05-05). eprint: <https://math.stackexchange.com/q/4693212>. URL: <https://math.stackexchange.com/q/4693212>.
- [8] TM Apostol. *Calculus, 2nd Edn., Vol. 2*. 1967.
- [9] Christopher M Bishop and Nasser M Nasrabadi. *Pattern Recognition and Machine Learning*. Vol. 4. 4. Springer, 2006.
- [10] Stephen Boyd, Stephen P. Boyd, and Lieven Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.
- [11] Gerd Brewka. *Artificial Intelligence—a Modern Approach by Stuart Russell and Peter Norvig, Prentice Hall. Series in Artificial Intelligence, Englewood Cliffs, NJ*. Vol. 11. Mar. 1996. URL: [https://www.cambridge.org/core/product/identifier/S0269888900007724/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0269888900007724/type/journal_article) (visited on 01/21/2023).
- [12] Robert Grover Brown and Patrick YC Hwang. *Introduction to Random Signals and Applied Kalman Filtering: With MATLAB Exercises and Solutions*. 1997.
- [13] Charles Casimiro. *Lecture notes in Statistical Signal Processing*. 2019.
- [14] Rama Chellappa and Sergios Theodoridis. *Signal Processing Theory and Machine Learning*. Academic Press, 2014. ISBN: 0-12-396502-0.
- [15] David Keun Cheng. *Field and Wave Electromagnetics*. Pearson Education India, 1989.

- [16] Jon Dattorro. *Convex Optimization & Euclidean Distance Geometry*. Lulu.com, 2010. ISBN: 0-615-19368-4.
- [17] Paulo SR Diniz. *Adaptive Filtering: Algorithms and Practical Implementation*. Nowell, MA: Kluwer Academic Publishers, 2002.
- [18] Paulo SR Diniz, Eduardo AB Da Silva, and Sergio L Netto. *Digital Signal Processing: System Analysis and Design*. Cambridge University Press, 2010. ISBN: 1-139-49157-1.
- [19] *Example Wikipedia Page*. URL: [https://en.wikipedia.org/wiki/Gauss%27s\\_law#Equation\\_involving\\_the\\_D\\_field](https://en.wikipedia.org/wiki/Gauss%27s_law#Equation_involving_the_D_field).
- [20] *Example Wikipedia Page*. URL: [https://en.wikipedia.org/wiki/Flux#Electric\\_flux](https://en.wikipedia.org/wiki/Flux#Electric_flux).
- [21] Andrea Goldsmith. *Wireless Communications*. Cambridge university press, 2005. ISBN: 0-521-83716-2.
- [22] Gene H Golub and Charles F Van Loan. *Matrix Computations*. JHU press, 2013. ISBN: 1-4214-0859-7.
- [23] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Illustrated edição. Cambridge, Massachusetts: The MIT Press, Nov. 18, 2016. ISBN: 978-0-262-03561-3.
- [24] Ronald L Graham et al. “Concrete Mathematics: A Foundation for Computer Science”. In: *Computers in Physics* 3.5 (1989), pp. 106–107. ISSN: 0894-1866.
- [25] Simon Haykin. *Neural Networks and Learning Machines, 3/E*. Pearson Education India, 2009. ISBN: 93-325-8625-X.
- [26] Simon S Haykin. *Adaptive Filter Theory*. Pearson Education India, 2002. ISBN: 81-317-0869-1.
- [27] Vinay K Ingle and John G Proakis. *Digital Signal Processing Using MATLAB*. Cole Publishing Company, 2000.
- [28] Steven M. Kay. *Fundamentals of Statistical Processing, Volume 2: Detection Theory*. Pearson Education India, 2009.
- [29] Basil Kouvaritakis and Mark Cannon. “Model Predictive Control”. In: *Switzerland: Springer International Publishing* 38 (2016).
- [30] Erwin Kreyszig, K Stroud, and G Stephenson. *Advanced Engineering Mathematics*. Vol. 9. John Wiley & Sons, Inc. 9 th edition, 2006 Page 2 of 6 Teaching methods ..., 2008.
- [31] Alberto Leon-Garcia. *Probability, Statistics, and Random Processes for Electrical Engineering*. 3rd ed. edição. Upper Saddle River, NJ: Prentice Hall, 2007. ISBN: 978-0-13-147122-1.
- [32] Josef Nosssek. *Adaptive and Array Signal Processing*. 2015.
- [33] Alan V. Oppenheim and Ronald W. Schaffer. *Discrete-Time Signal Processing: International Edition*. 3<sup>a</sup> edição. Upper Saddle River Munich: Pearson, Nov. 12, 2009. ISBN: 978-0-13-206709-6.

- [34] Kaare Brandt Petersen and Michael Syskind Pedersen. “The Matrix Cookbook”. In: *Technical University of Denmark* 7.15 (2008), p. 510.
- [35] John Proakis and Masoud Salehi. *Digital Communications*. 5th ed. edição. Boston: Mc Graw Hill, Jan. 1, 2007. ISBN: 978-0-07-295716-7.
- [36] Simon Ramo, John R Whinnery, and Theodore Van Duzer. *Fields and Waves in Communication Electronics*. John Wiley & Sons, 1994. ISBN: 81-265-1525-2.
- [37] Kenneth H Rosen. “Discrete Mathematics and Its Applications (7Th Edition)”. In: *William C Brown Pub* (2011).
- [38] Shayle R Searle and Andre I Khuri. *Matrix Algebra Useful for Statistics*. John Wiley & Sons, 2017. ISBN: 1-118-93514-4.
- [39] James Stewart. *Calculus*. Cengage Learning, 2011. ISBN: 1-133-17069-2.
- [40] Gilbert Strang et al. *Introduction to Linear Algebra*. Vol. 3. Wellesley-Cambridge Press Wellesley, MA, 1993.
- [41] Sergios Theodoridis. *Machine Learning: A Bayesian and Optimization Perspective*. 2nd ed. Academic Pr, 2020. ISBN: 978-0-12-818803-3.
- [42] Harry L Van Trees. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 2002. ISBN: 0-471-09390-4.
- [43] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons, 2004.