Notation

Rubem Vasconcelos Pacelli
rubem.engenharia@gmail.com

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

Version: July 24, 2024

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 . . . .

6 Communication systems
6.1 Common symbols . . . .
6.2 Fading multipath channels . . . .

7 Discrete mathematics
7.1 Quantifiers, inferences . . . .
7.2 Propositional Logic . . . .
7.3 Operations . . . .

8 Vector Calculus

9 Electromagnetic waves

10 Generic mathematical symbols

11 Abbreviations

1 Font notation

$a, b, c, \ldots, A, B, C, \ldots$	Scalars or tuples (the elements should be denoted in parentheses [44], although some authors also denote them in angle brackets [12])
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots$	Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$	Matrices
$\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{C}}, \ldots$	Tensors
$A, B, C, \ldots, \mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots, \mathbb{A}, \mathbb{B}, \mathbb{C}, \ldots$	Sets

2 Signals and functions

2.1 Time indexing

$x(t)$	Continuous-time t
$x[n], x[k], x[m], x[i], \ldots$ $x_n, x_k, x_m, x_i, \ldots$ $x(n), x(k), x(m), x(i), \ldots$	Discrete-time $n, k, m, i, \ldots$ (parenthesis should be adopted only if there are no continuous-time signals in the context to avoid ambiguity)
$x\left[((n-m))_N\right]$ [33], $x\left((n-m)\right)_N$ [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

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

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

2.4 Operations and symbols

$f : A \to B$	A function f whose domain is A and codomain is B
$\mathbf{f} : A \to \mathbb{R}^n$	A vector-valued function $\mathbf{f}$, i.e., $n \geq 2$
$f^n, x^n(t), x^n[k]$	nth power of the function f, $x[n]$ or $x(t)$
$f^{(n)}, x^{(n)}(t)$	nth 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., $f(\mathbf{x}) = \min\{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in \mathcal{A} \wedge (\mathbf{x}, \mathbf{y}) \in \mathrm{dom}\,(g)\}$, which is the greatest lower bound of this set [10, Appendix A.2.2]
$f(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{A}} g(\mathbf{x}, \mathbf{y})$	Supremum, i.e., $f(\mathbf{x}) = \max\{g(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in \mathcal{A} \wedge (\mathbf{x}, \mathbf{y}) \in \mathrm{dom}\,(g)\}$, which is the least upper bound of this set [10, Appendix A.2.2]
$f \circ g$	Composition of the functions f and g
$*$	Convolution (discrete or continuous)
$\circledast$ [18], Ⓝ [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 secod [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 ω only. However, in [27], the uppercase letters F and $\mathbf{\Omega}$ are used to denote the continuous-time frequency, while the lowercase f and ω denote the discrete-time frequency (Oppenheim [33] does not do it!)
Ω [27]	Continuous angular frequency (in rad/s), that is, $2\pi f$.
Ω_s	Sampling frequency (in rad/s), i.e., $2\pi f_s$
ω	Discrete angular frequency, i.e., ΩT_s [27, eq (3.27)]. As ω 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]
Ω_N [33], B	One-sided effective bandwidth of the continuous-time signal spectrum
ω_s [27]	Stop frequency
ω_p [27]	Pass frequency
$\Delta\omega$ [27]	$\omega_s - \omega_p$
ω_c [27]	Cutoff frequency
$s(t)$	Impulse train
$\mathrm{gdr}\left[H(e^{j\omega})\right]$ [33]	Group delay of $H(e^{j\omega})$
$\angle H(e^{j\omega})$ [33]	Phase response of $H(e^{j\omega})$
$\left\|H(e^{j\omega})\right\|$ [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

E $[\cdot]$, $\mathbf{E}[\cdot]$ [32], $E[\cdot]$, $\mathbb{E}[\cdot]$ [17]	Statistical expectation operator
$\cdot]$, $\mathbf{E}_u[\cdot]$ [32], $E_u[\cdot]$, $\mathbb{E}_u[\cdot]$	Statistical expectation operator with respect to u
$\langle\cdot\rangle$	Ensemble average
var $[\cdot]$ [32], VAR$[\cdot]$ [9, 26, 31, 37]	Variance operator
$\mathrm{var}_u[\cdot]$ $[\cdot]$, $\mathrm{VAR}_u[\cdot]$	Variance operator with respect to u
cov $[\cdot]$, COV$[\cdot]$	Covariance operator [9]
$\mathrm{cov}_u[\cdot]$, $\mathrm{COV}_u[\cdot]$	Covariance operator with respect to u
μ_x	Mean of the random variable x
$\boldsymbol{\mu}_{\mathbf{x}}, \mathbf{m}_{\mathbf{x}}$	Mean vector of the random variable $\mathbf{x}$ [13]
μ_n	nth-order moment of a random variable
σ_x^2, κ_2	Variance of the random variable x
$\mathcal{K}_x, \mu_4$	Kurtosis (4th-order moment) of the random variable x

κ_n	nth-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 (ω) frequency
$S_{x,y}(f), S_{x,y}(j\omega)$	Cross PSD of $x(t)$ and $y(t)$ in linear or angular (ω) frequency
$\mathbf{R}_\mathbf{x}$	(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]
$\mathbf{R}_\mathbf{xy}$	Cross-correlation matrix of $\mathbf{x}(n)$ and $\mathbf{y}(n)$
$\mathbf{r}_{\mathbf{x}d}$ [25], $\mathbf{p}_{\mathbf{x}d}$ [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]
$\mathbf{C}_\mathbf{x}, \mathbf{K}_\mathbf{x}, \mathbf{\Sigma}_\mathbf{x}, \mathrm{cov}\,[\mathbf{x}]$	(Auto)covariance matrix of $\mathbf{x}$ [9, 26, 31, 37, 45]
$\tilde{\mathbf{C}}_\mathbf{x}$[37]	Pseudocovariance matrix of $\mathbf{x}$
$c_{xy}(\tau), C_{xy}(\tau)$	Cross-covariance function of the signal $x(t)$ or $x[n]$ [32]
$\mathbf{C}_\mathbf{xy}, \mathbf{K}_\mathbf{xy}, \mathbf{\Sigma}_\mathbf{xy}$	Cross-covariance matrix of $\mathbf{x}$ and $\mathbf{y}$

3.3 Functions

$Q(\cdot)$	Q-function, i.e., $P\left[\mathcal{N}(0,1) > x\right]$ [37]
$\mathrm{erf}(\cdot)$	Error function [37]
$\mathrm{erfc}(\cdot)$	Complementary error function i.e., $\mathrm{erfc}(x) = 2Q(\sqrt{2}x) - \mathrm{erf}(x)$ [37]
$P[A]$	Probability of the event or set A [31]

$p(\cdot), f(\cdot)$	Probability density function (PDF) or probability mass function (PMF) [31]
$p(x \mid A)$	Conditional PDF or PMF [31]
$F(\cdot)$	Cumulative distribution function (CDF)
$\Phi_x(\omega), M_x(j\omega), E\left[e^{j\omega x}\right]$	First characteristic function (CF) of x [37, 44]
$M_x(t), \Phi_x(-jt), E\left[e^{tx}\right]$	Moment-generating function (MGF) of x [37, 44]
$\Psi_x(\omega), \ln \Phi_x(\omega), \ln E\left[e^{j\omega x}\right]$	Second characteristic function
$K_x(t), \ln E\left[e^{tx}\right], \ln M_x(t)$	Cumulant-generating function (CGF) of x [26]

3.4 Distributions

$\mathcal{N}(\mu, \sigma^2)$ [34]	Gaussian distribution of a random variable with mean μ and variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	Complex Gaussian distribution of a random variable with mean μ and variance σ^2
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution of a vector random variable with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Complex Gaussian distribution of a vector random variable with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathcal{U}(a, b)$	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 λ
$\Gamma(\alpha, \beta)$	Gamma distribution with shape parameter α and rate parameter β
$\Gamma(\alpha, \theta)$	Gamma distribution with shape parameter α and scale parameter $\theta = 1/\beta$
$\text{Nakagami}(m, \boldsymbol{\Omega})$	Nakagami-m distribution with shape parameter or fading figure m and spread, scale, or shape parameter $\boldsymbol{\Omega}$
$\text{Rayleigh}(\sigma)$	Rayleigh distribution with scale parameter σ

Rayleigh($\mathbf{\Omega}$)	Rayleigh distribution with the second moment $\mathbf{\Omega} = E\left[x^2\right] = 2\sigma^2$
Rice(s, σ)	Rice distribution with noncentrality parameter [37, p. 841] s and σ. s^2 represent the specular component power [37, p. 841]
Rice($\mathbf{\Omega}, K$), Rice(A, K)	Rice distribution with Rice factor $K = s^2/2\sigma^2$ and scale parameter $\mathbf{\Omega} = A = s^2 + 2\sigma^2 = 2\sigma^2(K+1)$ (Ω is prefered 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

$\mathbf{g}, \nabla f, \dfrac{\partial f}{\partial \mathbf{w}}$	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 ∇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
$\mathbf{g}_{\mathbf{x}}, \nabla_{\mathbf{w}} f, \dfrac{\partial f}{\partial \mathbf{w}}$	Gradient descent vector with respect $\mathbf{w}$ [9]
$\mathbf{J}, \dfrac{\partial \mathbf{y}^\top}{\partial \mathbf{x}}, \nabla \mathbf{y}^\top$ [25]	Jacobian matrix.
$\mathbf{H}, \dfrac{\partial^2 f}{\partial \mathbf{w}^2}, \nabla^2 f$ [25], $\nabla\nabla f$ [9]	Hessian matrix. The notation ∇^2 is sometimes used in Matrix Calculus to denote the second-order vector derivative. However, it must be avoided since, in Vector Calculus, ∇^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], $L(\mathbf{x})$ [14, 28]	Likelihood ratio function (also called likelihood ratio test (LRT) [28])
$\Lambda_l(\mathbf{x})$, $\mathcal{L}(\mathbf{x})$ [14], $l(\mathbf{x})$ [28]	Log-likelihood ratio (LLR [28]) function
$\hat{\rho}_{x,y}$	Estimated Pearson correlation coefficient between x and y
$\mathcal{R}_k$	kth Decision region
$x(t) \overset{m.s.e}{=} y(t)$	$x(t)$ equals $y(t)$ is the mean square error sense, that is $\mathrm{E}\left[\|x(t)-y(t)\|^2\right]=0$
$x(t) = \underset{N\to\infty}{\mathrm{l.i.m.}} \sum_{i=1}^{N} x_i \phi_i(t)$[46]	$\lim_{N\to\infty} \mathrm{E}\left[\left\|x(t)-\sum_{i=1}^{N} x_i\phi_i(t)\right\|^2\right]=0$ (l.i.m stands for "limit in the mean"). It is analogous to the $\overset{m.s.e}{=}$ notation, but denoting that they equal in the MSE sense only when $N \to \infty$

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

N	Number of instances/samples/examples[25], i.e., $n \in \{1, 2, \ldots, N\}$
N_{trn}	Number of instances in the training set, i.e., $n \in \{1, 2, \ldots, N_{\mathrm{trn}}\}$
N_{tst}	Number of instances in the test set, i.e., $n \in \{1, 2, \ldots, N_{\mathrm{tst}}\}$
N_{val}	Number of instances in the validation set, i.e., $n \in \{1, 2, \ldots, N_{\mathrm{val}}\}$
N_e	Number of epochs
N_a	Number os 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 lth 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 bais, which is usually ± 1. $+1$ is prefered [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 lth layer)
$y_{m_l}^{(l)}(n), \varphi\left(v_{m_l}^{(l)}(n)\right)$[25], $\mathbf{t}_{m_l}^{(l)}(n)$[9]	Output signal (target) of the m_lth neuron at the lth layer
$\mathbf{y}^{(l)}(n)$	Output signal of the lth 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 lth 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 lth 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 lth layer
$\mathbf{W}(n)$, $\begin{bmatrix} \mathbf{w}(1) & \mathbf{w}(2) & \cdots & \mathbf{w}(N) \end{bmatrix}^\top$	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 lth layer
$\mathbf{w}^{\star}, \mathbf{w}_o, \boldsymbol{\theta}^{\star}, \boldsymbol{\theta}_o$	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)
$\delta_{m_l}^{(l)}(n)$, $\frac{\partial \mathscr{E}(n)}{\partial v_{m_l}^{(l)}(n)}$	Local gradient of the m_lth neuron of the lth layer.
$\boldsymbol{\delta}^{(l)}(n)$	Vector of the local gradients of all neurons at the lth layer
$\mathbf{X}$, $\begin{bmatrix}\mathbf{x}(1) & \mathbf{x}(2) & \cdots & \mathbf{x}(N)\end{bmatrix}$	Data matrix [25]
$\eta(n)$	Learning rate hyperparameter [25]
$\mathscr{R}$	Bayes risk or average risk [25]
c_{ij}, C_{ij}	Misclassification cost in deciding in favor of class $\mathscr{C}_i$ (represented in the subspace $\mathscr{H}_i$) when the $\mathscr{C}_j$ is the true class (used in Bayes classifiers/detectors) [14, 25]
$\mathscr{C}_k$[25], $\mathcal{C}_k$ [9]	kth class
$\mathscr{T}$ [25], $\mathbb{X}$ [23]	Training set, i.e., the set $\{\mathbf{x}(n), d(n)\}$ that is used in the training phase.
$\mathscr{H}_k$	Subspace of the training vector belonging to the class $\mathscr{C}_k$
$\mathscr{H}$	Complete space of the input vector, i.e., $\mathscr{H}_1 \cup \mathscr{H}_2 \cup \cdots \mathscr{H}_K$
$\mathscr{X}$ [25]	Set of all vectors in the training, batch, validation, or test dataset that were misclassified
$\mathscr{E}(\mathbf{w}), \mathscr{E}(\mathbf{w}(n)), \mathscr{E}(n)$	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
$\Delta\mathscr{E}(\mathbf{w}(n)), \Delta\mathscr{E}(n), \mathscr{E}(\mathbf{w}(n+1)) - \mathscr{E}(\mathbf{w}(n))$	Cost function or objective function (the way it is written depends on the purpose of the text)
$\mathscr{E}_{\text{av}}(\cdot)$[25]	Error energy averaged over the training sample or the empirical risk

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

5 Linear Algebra

5.1 Common matrices and vectors

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

5.2 Indexing

$x_{i_1,i_2,\ldots,i_N}, [\mathcal{X}]_{i_1,i_2,\ldots,i_N}$	Element in the position $(i_1, i_2, \ldots, i_N)$ of the tensor $\mathcal{X}$
$\mathcal{X}^{(n)}$	nth tensor of a nontemporal sequence
$\mathbf{x}_n, \mathbf{x}_{:n}$	nth column of the matrix X
$\mathbf{x}_{n:}$	nth row of the matrix X
$\mathbf{x}_{i_1,\ldots,i_{n-1},:,i_{n+1},\ldots,i_N}$	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
$.^{\odot n}$	nth-order Hadamard power
$.^{\odot \frac{1}{n}}$	nth-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
$\mathbf{A}^\top, \mathbf{A}^T, \mathbf{A}^t, \mathbf{A}'$ [41]	Transpose
$\mathbf{A}^{-\top}$	Transpose of the inverse, i.e., $\left(\mathbf{A}^{-1}\right)^\top = (\mathbf{A}^\top)^{-1}$ [22, 36]
$\mathbf{A}^*$	Complex conjugate
$\mathbf{A}^{\mathsf{H}}$	Hermitian
$\\|\mathbf{A}\\|_{\mathrm{F}}$	Frobenius norm
$\\|\mathbf{A}\\|$	Matrix norm
$\lvert\mathbf{A}\rvert, \det(\mathbf{A})$	Determinant
$\operatorname{diag}(\mathbf{A})$	The elements in the diagonal of $\mathbf{A}$
$\operatorname{vec}[\mathbf{A}]$	Vectorization: stacks the columns of the matrix $\mathbf{A}$ into a long column vector

$\text{vec}_d\left[\mathbf{A}\right]$	Extracts the diagonal elements of a square matrix and returns them in a column vector
$\text{vec}_l\left[\mathbf{A}\right]$	Extracts the elements strictly below the main diagonal of a square matrix in a column-wise manner and returns them into a column vector
$\text{vec}_u\left[\mathbf{A}\right]$	Extracts the elements strictly above the main diagonal of a square matrix in a column-wise manner and returns them into a column vector
$\text{vec}_b\left[\mathbf{A}\right]$	Block vectorization operator: stacks square block matrices of the input into a long block column matrix
$\text{unvec}\left(\mathbf{A}\right)$	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_∞ norm, ∞-norm, or Chebyshev norm
$\text{diag}\left(\mathbf{a}\right)$	Diagonalization: a square, diagonal matrix with entries given by the vector $\mathbf{a}$

5.6 Decompositions

$\mathbf{\Lambda}$	Eigenvalue matrix [43]
$\mathbf{Q}$	Eigenvectors matrix; Orthogonal matrix of the QR decomposition[43]
$\mathbf{R}$	Upper triangular matrix of the QR decomposition[43]
$\mathbf{U}$	Left singular vectors[43]
$\mathbf{U}_r$	Left singular nondegenerated vectors
$\mathbf{\Sigma}$	Singular value matrix
$\mathbf{\Sigma}_r$	Singular value matrix with nonzero singular values in the main diagonal

$\mathbf{\Sigma}^{+}$	Singular value matrix of the pseudoinverse [43]
$\mathbf{\Sigma}_r^{+}$	Singular value matrix of the pseudoinverse with nonzero singular values in the main diagonal
$\mathbf{V}$	Right singular vectors [43]
$\mathbf{V}_r$	Right singular nondegenerated vectors
$\mathrm{eig}\,(\mathbf{A})$	Set of the eigenvalues of $\mathbf{A}$ [15, 31, 36]
$[\![\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots]\!]$	CANDECOMP/PARAFAC (CP) decomposition of the tensor $\mathcal{X}$ from the outer product of column vectors of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}, \ldots$
$[\![\boldsymbol{\lambda}; \mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots]\!]$	Normalized CANDECOMP/PARAFAC (CP) decomposition of the tensor $\mathcal{X}$ from the outer product of column vectors of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}, \ldots$

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
$\{1, 2, \ldots, n\}$	Discrete set containing the integer elements $1, 2, \ldots, n$
$\mathbb{B} = \{0, 1\}$ [35]	Boolean set
$\varnothing$	Empty set
$\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$	Real or complex space (field)

$\mathbb{K}^{I_1 \times I_2 \times \cdots \times I_N}$	$I_1 \times I_2 \times \cdots \times I_N$-dimensional real (or complex) space
$\mathbb{K}_{+}^{I_1 \times I_2 \times \cdots \times I_N}$ [10] [16, sec. 2.1.3]	Nonnegative real (or complex) orthant. The name orthant is the higher-dimensional generalization of the term *quadrant* from the classical Cartesian partition of $\mathbb{R}^2$ [16, sec 2.1.3]
$\mathbb{K}_{-}^{I_1 \times I_2 \times \cdots \times I_N}$ [10] [16, sec. 2.1.3]	Same, but for nonpositive real (or complex) orthant.
$\mathbb{K}_{++}^{I_1 \times I_2 \times \cdots \times I_N}$	Positive real (or complex) orthant, i.e., $\mathbb{K}_{++} = \mathbb{K}_{+} \setminus \{\mathbf{0}\}$ [10]
$\mathbb{K}_{--}^{I_1 \times I_2 \times \cdots \times I_N}$	Negative real (or complex) orthant, 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, sec. 2.2.2], $\mathcal{S}^n$ [10, sec. 1.6]	Conic set (see [10, p. 35]) of the symmetric matrices in $\mathbb{R}^{n \times n}$
$\mathbb{S}^{n\perp}$ [16, sec. 2.2.2]	Conic set of the skew-symmetric (also called antisymmetric) matrices in $\mathbb{R}^{n \times n}$
$\mathbb{S}_{+}^n$, $\mathcal{S}_{+}^n$ [10, sec. 2.2.5]	Conic set of the symmetric positive semidefinite matrices in $\mathbb{R}^{n \times n}$ [10]
$\mathbb{S}_{++}^n, \mathcal{S}_{++}^n$ [10, sec. 2.2.5]	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 $\mathbb{C}^{n \times n}$
conv A [10, p. 34]	Convex hull of the set A
aff A [10, p. 23]	Affine hull of the set A
∂A [16, sec. 2.1.7] bd A [10, appendix A.2]	Boundary of the set A
int A [16, sec. 2.1.6.1] [10, p. 2.1.3]	Interior of the set A
rel int A [16, sec. 2.1.6.1] relint A [10, p. 2.1.3]	Relative interior of the set A
cl A [10, appendix A.2] $\bar{A}$ [16, sec. 2.1.6.1]	Closure of A
$\mathcal{R}$ (?)	Ray
$\mathcal{H}$ (?)	Hyperplane
$\mathcal{H}_{+}, \mathcal{H}_{-}$ [16, sec. 2.4]	Positive/negative halfspace

$B(\mathbf{x}_c, r)$ [11, sec. 2.2.2]	Euclidean ball with radium r and centered at $\mathbf{x}_c$
$\mathcal{E}$ [11, sec. 2.2.2]	Ellipsoid
C [10, sec. 2.2.3]	Norm cone
K [11, sec. 2.4]	Proper cone
K^* [10, sec. 2.6]	Dual cone
$\mathcal{P}$ [10, sec. 2.2.4]	Polyhedra
S (?)	Simplex
C_α [10, sec. 3.1.6]	α-sublevel set
epi f [10, sec. 3.1.7]	Epigraph of the function f
hypo f [10, sec 3.1.7]	Hypograph of the function f

5.7.3 Spaces from matrices or vectors

$\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$	Vector space spanned by the argument vectors [22]
$\mathrm{C}(\mathbf{A})$, columnspace($\mathbf{A}$), range($\mathbf{A}$), $\operatorname{span}\{\mathbf{A}\}$, image($\mathbf{A}$)	Columnspace, range or image, i.e., the space $\operatorname{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$, where $\mathbf{a}_i$ is the ith column vector of the matrix $\mathbf{A}$ [32, 43]
$\mathrm{C}\left(\mathbf{A}^{\mathsf{H}}\right)$	Row space (also called left columnspace) [32, 43]
$\mathrm{N}(\mathbf{A})$, nullspace($\mathbf{A}$), null($\mathbf{A}$), kernel($\mathbf{A}$)	Nullspace (or kernel space) [32, 43, 44]
$\mathrm{N}\left(\mathbf{A}^{\mathsf{H}}\right)$	Left nullspace
$\operatorname{rank}\mathbf{A}$	Rank, that is, $\dim(\operatorname{span}\{\mathbf{A}\}) = \dim(\mathrm{C}(\mathbf{A}))$ [32]
nullity ($\mathbf{A}$)	Nullity of $\mathbf{A}$, i.e., $\dim(\mathrm{N}(\mathbf{A}))$

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 [40]

$A \cup B$	Set of union
$A \cap B$	Set of intersection
$A \times B$	Cartesian product
A^n	$\underbrace{A \times A \times \cdots \times A}_{n \text{ times}}$
$A^\perp$	Orthogonal complement of A, e.g., $\mathrm{N}(\mathbf{A}) = \mathrm{C}(\mathbf{A}^\top)^\perp$ [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 $\mathbf{v} \in \{\sum \mathbf{a}_i \mid \mathbf{a}_i \in S_i, i = 1, \ldots, k\}$ 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., $\mathrm{C}(\mathbf{A}^\top) \overset{\perp}{\oplus} \mathrm{C}(\mathbf{A}^\top)^\perp = \mathbb{R}^n$ (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

$\mathcal{X} \leq 0$	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 baseband signal, in Hz
W	One-sided bandwidth of the baseband 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[37] 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
ϕ	Signal phase
ϕ_0	Initial phase
η_{RF}, w_{RF}	Noise in RF
η_{FI}, w_{FI}	Noise in FI
η, w	Noise in baseband
τ	Timing delay
$\Delta\tau$	Timing error (delay - estimated)
φ	Phase offset
$\Delta\varphi$	Phase error (offset - estimated)
f_d	Linear Doppler frequency
Δf_d	Frequency error (Doppler frequency - estimated)
ν	Angular Doppler frequency
$\Delta\nu$	Frequency error (Doppler frequency - estimated)
γ, A	Transmitted signal amplitude
γ_0, A_0	Combined effect of the path loss and antenna gain

6.2 Fading multipath channels

$t \overset{\mathcal{F}}{\leftrightarrow} \lambda$ [37]	Support temporal of the signal. λ is obtained after taking the Fourier transform on t.
$\tau \overset{\mathcal{F}}{\leftrightarrow} f$ [37]	Second support temporal of the signal ($c(t)$ varies with with the input at the time τ). f is obtained after taking the Fourier transform on τ.
$c(t,\tau)$ [37]	Complex envelope of the channel response at the time t due to an impulse applied at the $t-\tau$
$C(f,t)$ [37]	Transfer function of $c(t,\tau)$ in τ

$\alpha(t,\tau)$ [37]	Attenuation of $c(t,\tau)$, i.e., $c(t,\tau) = \alpha(t,\tau)e^{e\pi f_c \tau}$
$R_c(\tau_1,\tau_2,\Delta t)$ [37]	Autocorrelation function of $c(t,\tau)$, i.e., $R_c(\tau_1,\tau_2,\Delta t) = \mathrm{E}\left[c^*(t,\tau_1), c^*(t+\Delta t,\tau_2)\right]$
$R_c(\tau,\Delta t)$ [37]	Autocorrelation function of $c(t,\tau)$ assuming uncorrelated scattering
$R_c(\tau)$, $R_c(\tau,\Delta t)\big\|_{\Delta t=0}$ [37]	Multipath intensity profile or delay power spectrum
$R_C(\Delta f,\Delta t), R_C(f_1,f_2;\Delta t)$, $\mathrm{E}\left[C(f_1,t), C(f_2,t+\Delta t)\right]$, $\mathcal{F}_\tau\left\{R_c(\tau,\Delta t)\right\}$ [21]	Spaced-frequency, spaced-time correlation function ($\Delta f = f_2 - f_1$)
$R_C(\Delta f)$, $R_C(\Delta f,\Delta t)\big\|_{\Delta t=0}$ [37], $\mathcal{F}\left\{R_c(\tau)\right\}$ [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 [37]
T_m	Multipath spread of the channel, that is, the time interval in which $R_c(\tau)$ is nonzero ($T_m \approx 1/(\Delta f)_c$) [37]
$R_C(\Delta t)$, $R_C(\Delta f,\Delta t)\big\|_{\Delta f=0}$	Spaced-time correlation function [37]
$S_C(\lambda)$ [37], $\mathcal{F}\left\{R_C(\Delta t)\right\}$ [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 [37]
B_m	Multipath spread of the channel, that is, the frequency interval in which $S_c(\lambda)$ is nonzero ($B_d \approx 1/(\Delta t)_c$) [37]
$S_C(\tau,\lambda)$ [37], $\mathcal{F}_{\Delta f,\Delta t}\left\{R_C(\Delta f,\Delta t)\right\}$ [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 [40]
$\in$	Belongs to [24]

$\notin$	Does not belong to [24]
$\because$	Because [24]
$\mid, :$	Such that, sometimes that parentheses is used [24]
$, , (\cdot)$	Used to separate the quantifier with restricted domain from its scope, e.g., $\forall \ \ x < 0 \left(x^2 > 0\right)$ or $\forall \ \ x < 0, x^2 > 0$ [24]
$\therefore$	Therefore [24]

7.2 Propositional Logic

$\neg a$	Logical negation of a [40]
$a \wedge b$	Conjunction (logical AND) operator between a and b[40]
$a \vee b$	Disjunction (logical OR) operator between a and b[40]
$a \oplus b$	Exclusive OR (logical XOR) operator between a and b[40]
$a \rightarrow b$	Implication (or conditional) statement[40]
$a \leftrightarrow b$	Bi-implication (or biconditional) statement, i.e., $(a \rightarrow b) \wedge (b \ \rightarrow \ a)$ [40]
$a \equiv b, a \iff b, a \Leftrightarrow b$	Logical equivalence, i.e., $a \leftrightarrow b$ is a tautology[40]

7.3 Operations

$\lvert a \rvert$	Absolute value of a
$\log$	Base-10 logarithm or decimal logarithm
$\ln$	Natual logarithm
$\mathrm{Re}\left\{x\right\}$	Real part of x
$\mathrm{Im}\left\{x\right\}$	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 [40]
$x \equiv y \pmod m$	Congruent, i.e., $m\backslash(x - y)$ [40]
$\mathrm{frac}\left(x\right)$	Fractional part, i.e., $x \bmod 1$ [24]
$a\backslash b$ [24, Section 4.1], $a \mid b$ [40]	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], $a \nmid b$ [40]	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

∇f[42], $\mathrm{grad} f$[38]	Vector differential operator (Nabla symbol), i.e., ∇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[42]
$\mathbf{l}(x, y, z)$ [38], $\mathbf{r}(x, y, z)$ [42], $x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$	Vector position, i.e., (x, y, z).
$\mathbf{l}(t)$	Vector position parametrized by t, i.e., $(x(t), y(t), z(t))$ [38, 42]
$\mathbf{l}'(t), \mathrm{d}\mathbf{l}\,/\mathrm{d}t$	First derivative of $\mathbf{l}(t)$, i.e., the tangent vector of the curve $(x(t), y(t), z(t))$ [42]
$\mathbf{u}(t)$[30] $\mathbf{T}(t)$[42], $\mathrm{d}\mathbf{l}\,(t)$[38]	Tangent unit vector of $\mathbf{l}(t)$, i.e., $\mathbf{u}(t) = \mathbf{l}'(t)/\lvert\mathbf{l}'(t)\rvert$
$\mathbf{n}(t), \left(\frac{y'(t)}{\lvert\mathbf{l}'(t)\rvert}, -\frac{x'(t)}{\lvert\mathbf{l}'(t)\rvert}\right)$	Normal vector of $\mathbf{l}(t)$, i.e., $\mathbf{n}(t) \perp \mathbf{T}(t)$[42]
C	Contour that traveled by $\mathbf{l}(t)$, for $a \leq t \leq b$ [42]
$L, L(C)$	Total length of the contour C (which can be defined the vector $\mathbf{l}$, parametrized by t), i.e., $L_C = \int_a^b \lvert\mathbf{l}'(t)\rvert\,\mathrm{d}t$[42]
$s(t)$	Length of the arc, which can be defined by the vector $\mathbf{l}$ and t, that is, $s(t) = \int_a^t \lvert\mathbf{l}'(u)\rvert\,\mathrm{d}u$ $(s(b) = L)$[42]
$\mathrm{d}s$	Differential operator of the length of the contour C, i.e., $\mathrm{d}s = \lvert\mathbf{l}'(t)\rvert\,\mathrm{d}t$ [42]
$\int_C f(\mathbf{l})\,\mathrm{d}s$, $\int_a^b f(\mathbf{l}(t))\lvert\mathbf{l}'(t)\rvert\,\mathrm{d}t$	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”
θ [38]	Angle between the contour C and the vector field $\mathbf{F}$

$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{l}$, $\int_a^b \mathbf{F}(\mathbf{l}(t)) \cdot \mathbf{l}'(t)\,\mathrm{d}t$ [8, 42], $\int_C \mathbf{F} \cdot \mathbf{u}\,\mathrm{d}s$, $\int_C \mathbf{F} \cos(\theta)\,\mathrm{d}s$ [38]	Line integral of vector field $\mathbf{F}$ along the contour C
$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{u}$ [38]	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
$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F}$, $\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{F} \cdot \mathrm{d}\mathbf{l}$	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]
$\oint_C$, $\ointctrclockwise_C$	Line integral along the closed contour C. The arrow indicates the contour integral orientation, which is counterclockwise, by default. In the context of integrals in the complex plane, it is also called "closed contour integral".
$\oiint_S$	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)
$\mathbf{l}_u$	$(\partial x/\partial u, \partial y/\partial u, \partial z/\partial u)$
$\mathbf{l}_v$	$(\partial x/\partial v, \partial y/\partial v, \partial z/\partial v)$
$\mathrm{d}A$	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) [42]
D, R	Integration domain in which $\mathrm{d}A$ is integrated, i.e., $\iint_D f\,\mathrm{d}A$. R is preferred when the integration domain is a rectangle, while D is used when it has nonrectangular shape [42]
S	Smooth surface $S \subset \mathbb{R}^3$, i.e., a 2D area in a 3D space
$\mathrm{d}S$, $\lvert\mathbf{l}_u \times \mathbf{l}_v\rvert\,\mathrm{d}A$	Differential operator of a 2D area in a 3D domain (an surface). Note that $\mathrm{d}S = \lvert\mathbf{l}_u \times \mathbf{l}_v\rvert\,\mathrm{d}A$ should be accompanied with the change of the integration interval(from S to D)

$A(S)$, $\iint_S \mathrm{d}S$, $\iint_D \lvert \mathbf{l}_u \times \mathbf{l}_v \rvert \,\mathrm{d}A$	Area of the surface S parametrized by (u, v), in which $\mathrm{d}A$ is the area defined in the D domain (which is form by the u-by-v graph)
$\mathrm{d}V$	Differential operator of a shape volume (denoted by E) in $\mathbb{R}^3$ domain, i.e., $\iiint_E \mathrm{d}V = V$
E	Integration domain in which $\mathrm{d}V$ is integrated, i.e., $\iiint_E f \,\mathrm{d}V$ [42]
V, $\iint_D f \,\mathrm{d}A$, $\iiint_E f \,\mathrm{d}V$	Volume of the function f over the regions D (in the case of double integrals) or E (in the case of triple integrals)
$\iint_S f \,\mathrm{d}S$, $\iint_D f \lvert \mathbf{l}_u \times \mathbf{l}_v \rvert \,\mathrm{d}A$	Surface integral over S
$\mathbf{n}(u, v)$, $\frac{\mathbf{l}_u(u,v) \times \mathbf{l}_v(u,v)}{\lvert \mathbf{l}_u(u,v) \times \mathbf{l}_v(u,v) \rvert}$	Normal vector of of the smooth surface S
$\iint_S \mathbf{F} \cdot \mathbf{n} \,\mathrm{d}S$, $\iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$, $\iint_D \mathbf{F} \cdot (\mathbf{l}_u \times \mathbf{l}_v) \,\mathrm{d}A$	Flux integral of vector field $\mathbf{F}$ through the smooth surface S ($\mathbf{n}\,\mathrm{d}S \triangleq \mathrm{d}\mathbf{S}$)
$\oiint_S \mathbf{F} \cdot \mathbf{n} \,\mathrm{d}S$, $\oiint_S \mathbf{F} \cdot \mathrm{d}\mathbf{S}$, $\iint_D \mathbf{F} \cdot (\mathbf{l}_u \times \mathbf{l}_v) \,\mathrm{d}A$	Flux integral of vector field $\mathbf{F}$ through the smooth and closed surface S ($\mathbf{n}\,\mathrm{d}S \triangleq \mathrm{d}\mathbf{S}$)
$\nabla \times \mathbf{F}$, curl $\mathbf{F}$	Curl (rotacional) of the vector field $\mathbf{F}$
$\nabla \cdot \mathbf{F}$, div $\mathbf{F}$	Divercence 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 $f : \mathbb{R}^n \to \mathbb{R}$)
$\nabla^2 \mathbf{F}, \nabla \times \nabla \times \mathbf{F} - \nabla(\nabla \cdot \mathbf{F}), \Delta \mathbf{F}$, $(\partial^2 \mathbf{F}/\partial x^2, \partial^2 \mathbf{F}/\partial y^2, \partial^2 \mathbf{F}/\partial z^2)$	Vector Laplacian operator (performed on a vector field, i.e., a vector-valued function, $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$). ∇^2 denotes the scalar (vector) Laplacian if the function is scalar-valued (vector-valued). The notation Δ must be avoided as it is overused in many contexts

9 Electromagnetic waves

$\boldsymbol{\Phi}$	Electric flux (scalar) (in V m)
$\mathbf{H}$	Magnetic field vector (in A/m)
$\mathbf{B}$	Magnetic flux density vector (in $\mathrm{Wb/m^2 = T}$)
Φ[**chengFieldWaveElectromagnetics1989**]	Magnetic flux

q_f, q_free, Q_free[19]	Free electric charge (in C)
q_b, q_bound, Q_bound[19]	Bound electric charge (in C)
q, $q_\mathrm{f} + q_\mathrm{b}$	Electric charge (in C)
ρ_f[1],ρ_free [19]	Free electric charge density
ρ_b[1], ρ_bound [19]	Electric charge density
$\rho, \rho_\mathrm{f} + \rho_\mathrm{b}$	Electric charge density (it can be in $\mathrm{C/m^3}, \mathrm{C/m^2}$ or C/m depending whether it is a volume, surface, or line shapes)
$\mathbf{f}$[38], $\mathbf{F}$[2]	Electrostatic force (Coulomb force), (in $\mathrm{kg\,m/s^2}$).
ε	Electric permittivity(in F/m). If the medium is isotropic, it is a scalar. If it is anisotropic, it is a tensor. [38]
ε_r	Relative electric permittivity or dielectric constant (in F/m) [38]
ε_0	Electric permittivity in vacuum, $8.854 \times 10^{-12}\,\mathrm{F/m}$ [38]
$\mathbf{E}$	Electric field vector (in V/m)
σ	Electric conductivity (in S/m)
$\mathbf{J}$, $\mathbf{J}_e$[39]	Electric current density vector (in $\mathrm{A/m^2}$)
$\mathbf{J}_m$[**chengFieldWaveElectromagnetics1989**]	Magnetization current density vector (in $\mathrm{A/m^2}$)
$\mathbf{D}$	Electric flux density, electric displacement, or electric induction vector (in $\mathrm{C/m^2}$)
U	Electric potential energy
V[**chengFieldWaveElectromagnetics1989**, 3], $\mathbf{\Phi}$[38]	Electric potential (in voltage, V). However, keep in mind that there is a subtle difference between both definitions [4]
$\mathbf{\Phi}_E$[20], $\oiint_S \mathbf{E}\,\mathrm{d}\mathbf{S}$	Electric flux (in V m)
$\mathbf{\Phi}_D$[19], Ψ[38], $\oiint_S \mathbf{D}\,\mathrm{d}\mathbf{S}$	Electric flux (**D**-field flux)
$\mathbf{P}$	Electric polarization of the material (in $\mathrm{C/m^2}$)
χ_e	Electric susceptibility (for linear and isotropic materials)
μ	Magnetic permeability
μ_0	Magnetic permeability in vacuum

10 Generic mathematical symbols

■	Q.E.D.
$\triangleq$	Equal by definition
$:=, \leftarrow$	Assignment [40]
$\neq$	Not equal
∞	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.

[2] URL: https://en.wikipedia.org/wiki/Coulomb%27s_law.

[3] URL: 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 ∇^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 ∇^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. ISBN: 978-0-387-31073-2.

[10] Stephen Boyd, Stephen P. Boyd, and Lieven Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.

[11] Stephen Boyd and Lieven Vandenberghe. "Additional Exercises for Convex Optimization". In: ().

[12] 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 (visited on 01/21/2023).

[13] Robert Grover Brown and Patrick YC Hwang. *Introduction to Random Signals and Applied Kalman Filtering: With MATLAB Exercises and Solutions*. 1997.

[14] Charles Casimiro. *Lecture notes in Statistical Signal Processing*. 2019.

[15] Rama Chellappa and Sergios Theodoridis. *Signal Processing Theory and Machine Learning*. Academic Press, 2014. ISBN: 0-12-396502-0.

[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.

[20] *Example Wikipedia Page*. URL: 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. Upper Saddle River, NJ: Prentice Hall, 2007. ISBN: 978-0-13-147122-1.

[32] Josef Nossek. *Adaptive and Array Signal Processing*. 2015.

[33] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing: International Edition*. 3ª edição. Upper Saddle River Munich: Pearson, Nov. 12, 2009. ISBN: 978-0-13-206709-6.

[34] Athanasios Papoulis et al. *Probability, Random Variables and Stochastic Processes*. 4th ed. edição. Boston: McGraw-Hill Education, 2001. ISBN: 978-0-07-366011-0.

[35] Ian Parberry. *Circuit Complexity and Neural Networks*. MIT press, 1994. URL: https://books.google.com.br/books?hl=en&lr=&id=Mc3sIKru91wC&oi=fnd&pg=PR9&dq=Circuit+Complexity+and+Neural+Networks&ots=jJJyyZf-Ie&sig=pXP7VlnDM1PWlJs17kvIZ3UFqD0 (visited on 04/22/2024).

[36] Kaare Brandt Petersen and Michael Syskind Pedersen. "The Matrix Cookbook". In: *Technical University of Denmark* 7.15 (2008), p. 510.

[37] John Proakis and Masoud Salehi. *Digital Communications*. 5th ed. Boston: Mc Graw Hill, Jan. 1, 2007. ISBN: 978-0-07-295716-7.

[38] Simon Ramo, John R. Whinnery, and Theodore Van Duzer. *Fields and Waves in Communication Electronics*. John Wiley & Sons, 1994.

[39] Charles Rino. *The theory of scintillation with applications in remote sensing*. John Wiley & Sons, 2011.

[40] Kenneth H Rosen. "Discrete Mathematics and Its Applications (7Th Editio)". In: *William C Brown Pub* (2011).

[41] Shayle R Searle and Andre I Khuri. *Matrix Algebra Useful for Statistics*. John Wiley & Sons, 2017. ISBN: 1-118-93514-4.

[42] James Stewart. *Calculus*. Cengage Learning, 2011. ISBN: 1-133-17069-2.

[43] Gilbert Strang et al. *Introduction to Linear Algebra*. Vol. 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[44] Sergios Theodoridis. *Machine Learning: A Bayesian and Optimization Perspective*. 2nd ed. Academic Pr, 2020. ISBN: 978-0-12-818803-3.

[45] Harry L Van Trees. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. John Wiley & Sons, 2002. ISBN: 0-471-09390-4.

[46] Harry L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. John Wiley & Sons, 2004. ISBN: 978-0-471-22108-1. DOI: 10.1002/0471221082.