Appendix B

Notational conventions

B.1 List of symbols

N	dimension of feature space	L	primal Lagrangian
$y \in Y$	output and output space	W	dual Lagrangian
$\mathbf{x} \in X$	input and input space	$\|\|\cdot\|\|_p$	p-norm, default is 2-norm
$\|\|\mathbf{A}\|\|_F$	Frobenius norm of a matrix	$\|\|\mathbf{A}\|\|$	spectral/2-norm of a matrix
F	feature space	ln	natural logarithm
$\mathcal{F}$	class of real-valued functions	e	base of the natural log
$\mathcal{L}$	class of linear functions	log	log to the base 2
$\langle \mathbf{x}, \mathbf{z} \rangle$	inner product of $\mathbf{x}$ and $\mathbf{z}$	$\mathbf{x}'$, $\mathbf{X}'$	transpose of vector, matrix
$\boldsymbol{\phi}$	mapping to feature space	$\mathbb{N}$, $\mathbb{R}$	natural, real numbers
$\kappa(\mathbf{x}, \mathbf{z})$	kernel $\langle \boldsymbol{\phi}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{z}) \rangle$	S	training set
$f(\mathbf{x})$	real-valued function	ℓ	training set size
n	dimension of input space	$\boldsymbol{\phi}(S)$	training set in feature space
R	radius containing the data	η	learning rate
$\mathcal{H}$	Heaviside function	ε	error probability
$\mathbf{w}$	weight vector	δ	confidence
b	bias	γ	margin
$\boldsymbol{\alpha}$	dual variables	ξ	slack variables
$\mathbf{C}$	covariance matrix	$\mathbf{I}$	identity matrix
$(x)_+$	equals x, if $x \geq 0$ else 0	$\mathbf{K}$	kernel matrix
$\operatorname{sgn}(x)$	equals 1, if $x \geq 0$ else -1	#	cardinality of a set

B.2 Notation for Tables

Definition B.1 [Kernel matrix displays] We use a standard notation for displaying kernel matrices as

K	1	2	$\cdots$	ℓ
1	$\kappa(\mathbf{x}_1, \mathbf{x}_1)$	$\kappa(\mathbf{x}_1, \mathbf{x}_2)$	$\cdots$	$\kappa(\mathbf{x}_1, \mathbf{x}_\ell)$
2	$\kappa(\mathbf{x}_2, \mathbf{x}_1)$	$\kappa(\mathbf{x}_2, \mathbf{x}_2)$	$\cdots$	$\kappa(\mathbf{x}_2, \mathbf{x}_\ell)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
ℓ	$\kappa(\mathbf{x}_\ell, \mathbf{x}_1)$	$\kappa(\mathbf{x}_\ell, \mathbf{x}_2)$	$\cdots$	$\kappa(\mathbf{x}_\ell, \mathbf{x}_\ell)$

where the symbol **K** in the top right corner indicates that the table represents a kernel matrix. ■

Definition B.2 [Dynamic programming tables] Dynamic programming tables are displayed in a table with first row and column used for indices and the top left cell marked with DP, as, for example, in the ANOVA dynamic programming table:

DP	1	2	$\cdots$	n
0	1	1	$\cdots$	1
1	$x_1 z_1$	$x_1 z_1 + x_2 z_2$	$\cdots$	$\sum_{i=1}^{n} x_i z_i$
2	0	$\kappa_2^2(\mathbf{x}, \mathbf{z})$	$\cdots$	$\kappa_2^n(\mathbf{x}, \mathbf{z})$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
d	0	0	$\cdots$	$\kappa_d^n(\mathbf{x}, \mathbf{z})$

■