

# Contents

<b>1</b>	<b>About</b>	<b>1</b>
<b>2</b>	<b>Notation</b>	<b>5</b>
<b>3</b>	<b>Linear Algebra</b>	<b>6</b>
3.1	Vector spaces . . . . .	6
3.1.1	Euclidean space . . . . .	6
3.1.2	Subspaces . . . . .	7
3.2	Linear maps . . . . .	7
3.2.1	The matrix of a linear map . . . . .	8
3.2.2	Nullspace, range . . . . .	9
3.3	Metric spaces . . . . .	9
3.4	Normed spaces . . . . .	9
3.5	Inner product spaces . . . . .	10
3.5.1	Pythagorean Theorem . . . . .	11
3.5.2	Cauchy-Schwarz inequality . . . . .	11
3.5.3	Orthogonal complements and projections . . . . .	12
3.6	Eigenthings . . . . .	15
3.7	Trace . . . . .	15
3.8	Determinant . . . . .	16
3.9	Orthogonal matrices . . . . .	16
3.10	Symmetric matrices . . . . .	17
3.10.1	Rayleigh quotients . . . . .	17
3.11	Positive (semi-)definite matrices . . . . .	18
3.11.1	The geometry of positive definite quadratic forms . . . . .	19
3.12	Singular value decomposition . . . . .	20
3.13	Fundamental Theorem of Linear Algebra . . . . .	21
3.14	Operator and matrix norms . . . . .	22
3.15	Low-rank approximation . . . . .	24
3.16	Pseudoinverses . . . . .	25
3.17	Some useful matrix identities . . . . .	26
3.17.1	Matrix-vector product as linear combination of matrix columns . . . . .	26
3.17.2	Sum of outer products as matrix-matrix product . . . . .	26
3.17.3	Quadratic forms . . . . .	26
<b>4</b>	<b>Calculus and Optimization</b>	<b>27</b>

4.1	Extrema . . . . .	27
4.2	Gradients . . . . .	27
4.3	The Jacobian . . . . .	27
4.4	The Hessian . . . . .	28
4.5	Matrix calculus . . . . .	28
4.5.1	The chain rule . . . . .	28
4.6	Taylor's theorem . . . . .	29
4.7	Conditions for local minima . . . . .	29
4.8	Convexity . . . . .	31
4.8.1	Convex sets . . . . .	31
4.8.2	Basics of convex functions . . . . .	31
4.8.3	Consequences of convexity . . . . .	32
4.8.4	Showing that a function is convex . . . . .	33
4.8.5	Examples . . . . .	36
<b>5</b>	<b>Probability</b>	<b>37</b>
5.1	Basics . . . . .	37
5.1.1	Conditional probability . . . . .	38
5.1.2	Chain rule . . . . .	38
5.1.3	Bayes' rule . . . . .	38
5.2	Random variables . . . . .	39
5.2.1	The cumulative distribution function . . . . .	39
5.2.2	Discrete random variables . . . . .	40
5.2.3	Continuous random variables . . . . .	40
5.2.4	Other kinds of random variables . . . . .	40
5.3	Joint distributions . . . . .	41
5.3.1	Independence of random variables . . . . .	41
5.3.2	Marginal distributions . . . . .	41
5.4	Great Expectations . . . . .	41
5.4.1	Properties of expected value . . . . .	42
5.5	Variance . . . . .	42
5.5.1	Properties of variance . . . . .	42
5.5.2	Standard deviation . . . . .	42
5.6	Covariance . . . . .	43
5.6.1	Correlation . . . . .	43
5.7	Random vectors . . . . .	43

5.8	Estimation of Parameters . . . . .	44
5.8.1	Maximum likelihood estimation . . . . .	44
5.8.2	Maximum a posteriori estimation . . . . .	45
5.9	The Gaussian distribution . . . . .	45
5.9.1	The geometry of multivariate Gaussians . . . . .	45
<b>References</b>		<b>47</b>