

# Deep Learning

Ian Goodfellow  
Yoshua Bengio  
Aaron Courville

# Contents

<b>Website</b>	<b>vii</b>
<b>Acknowledgments</b>	<b>viii</b>
<b>Notation</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Who Should Read This Book? . . . . .	8
1.2 Historical Trends in Deep Learning . . . . .	11
<b>I Applied Math and Machine Learning Basics</b>	<b>29</b>
<b>2 Linear Algebra</b>	<b>31</b>
2.1 Scalars, Vectors, Matrices and Tensors . . . . .	31
2.2 Multiplying Matrices and Vectors . . . . .	34
2.3 Identity and Inverse Matrices . . . . .	36
2.4 Linear Dependence and Span . . . . .	37
2.5 Norms . . . . .	39
2.6 Special Kinds of Matrices and Vectors . . . . .	40
2.7 Eigendecomposition . . . . .	42
2.8 Singular Value Decomposition . . . . .	44
2.9 The Moore-Penrose Pseudoinverse . . . . .	45
2.10 The Trace Operator . . . . .	46
2.11 The Determinant . . . . .	47
2.12 Example: Principal Components Analysis . . . . .	48
<b>3 Probability and Information Theory</b>	<b>53</b>
3.1 Why Probability? . . . . .	54











