

13.5	Manifold Interpretation of PCA . . . . .	499
<b>14</b>	<b>Autoencoders</b>	<b>502</b>
14.1	Undercomplete Autoencoders . . . . .	503
14.2	Regularized Autoencoders . . . . .	504
14.3	Representational Power, Layer Size and Depth . . . . .	508
14.4	Stochastic Encoders and Decoders . . . . .	509
14.5	Denoising Autoencoders . . . . .	510
14.6	Learning Manifolds with Autoencoders . . . . .	515
14.7	Contractive Autoencoders . . . . .	521
14.8	Predictive Sparse Decomposition . . . . .	523
14.9	Applications of Autoencoders . . . . .	524
<b>15</b>	<b>Representation Learning</b>	<b>526</b>
15.1	Greedy Layer-Wise Unsupervised Pretraining . . . . .	528
15.2	Transfer Learning and Domain Adaptation . . . . .	536
15.3	Semi-Supervised Disentangling of Causal Factors . . . . .	541
15.4	Distributed Representation . . . . .	546
15.5	Exponential Gains from Depth . . . . .	553
15.6	Providing Clues to Discover Underlying Causes . . . . .	554
<b>16</b>	<b>Structured Probabilistic Models for Deep Learning</b>	<b>558</b>
16.1	The Challenge of Unstructured Modeling . . . . .	559
16.2	Using Graphs to Describe Model Structure . . . . .	563
16.3	Sampling from Graphical Models . . . . .	580
16.4	Advantages of Structured Modeling . . . . .	582
16.5	Learning about Dependencies . . . . .	582
16.6	Inference and Approximate Inference . . . . .	584
16.7	The Deep Learning Approach to Structured Probabilistic Models	585
<b>17</b>	<b>Monte Carlo Methods</b>	<b>590</b>
17.1	Sampling and Monte Carlo Methods . . . . .	590
17.2	Importance Sampling . . . . .	592
17.3	Markov Chain Monte Carlo Methods . . . . .	595
17.4	Gibbs Sampling . . . . .	599
17.5	The Challenge of Mixing between Separated Modes . . . . .	599
<b>18</b>	<b>Confronting the Partition Function</b>	<b>605</b>
18.1	The Log-Likelihood Gradient . . . . .	606
18.2	Stochastic Maximum Likelihood and Contrastive Divergence . . .	607