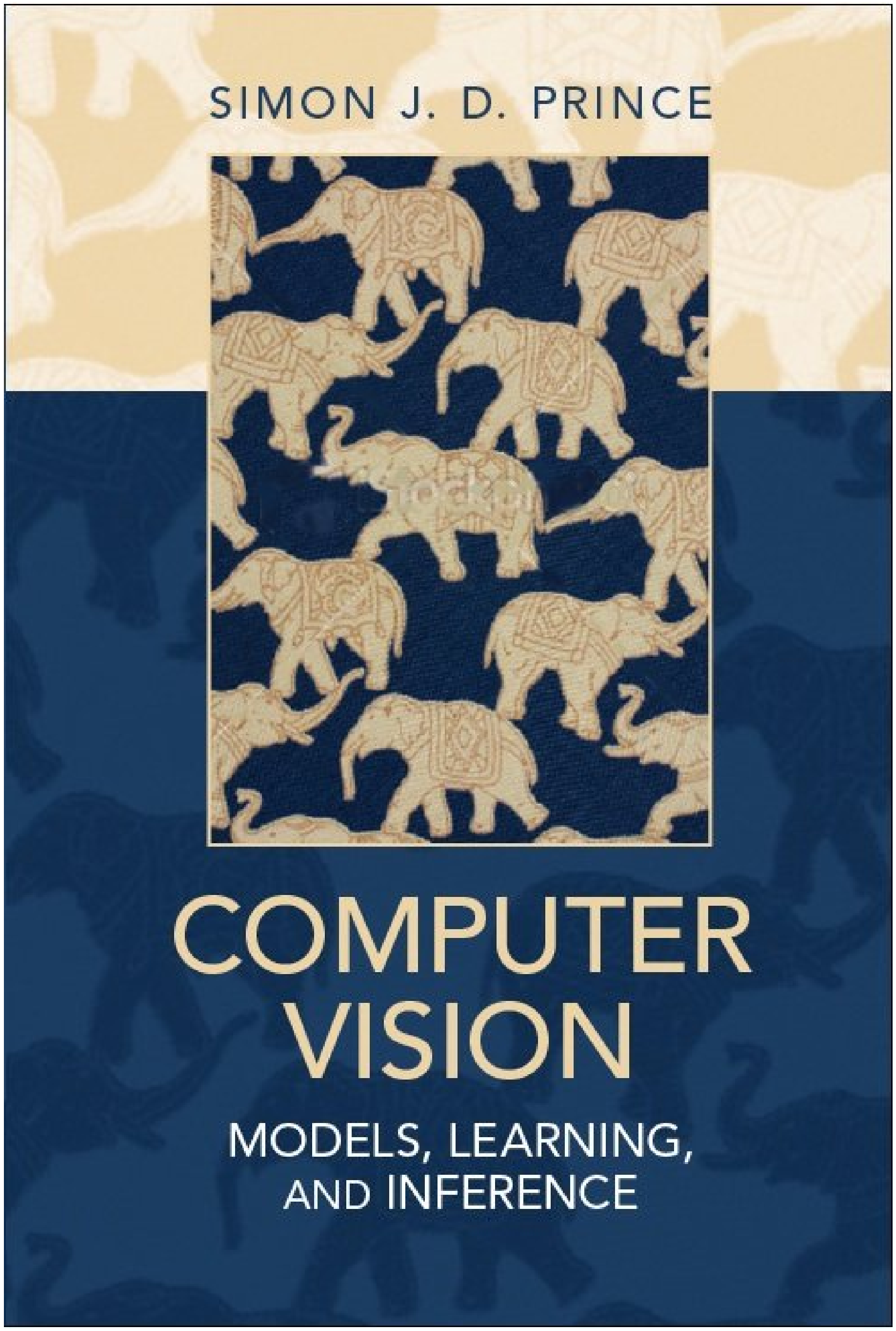

Errata

June 10, 2015

Errata

This document accompanies the book “Computer vision: models, learning, and inference” by Simon J.D. Prince, (http://www.computervisionmodels.com). It contains errata from the first and second printed editions (June 2012 and October 2014). Errata marked in red are in the first edition only. Errata marked in blue are in both printings.

f you think you find an error or ambiguity that is not covered in this booklet, then please mail me at s.prince@cs.ucl.ac.uk. It doesn’t matter how small or unimportant it seems and it doesn’t matter if you are not sure. I’d still like to hear about it! I’ll make sure that the next printing is correct and that your name is added to the acknowledgements at the start of the book.

Thanks go to Dan Farmer, Samit Shah, Martijn van der Veen, Gabriel Brostow, Marco Brambilla, Sebastian Stabinger, Tamaki Toru, Stefan Stavref, Xiaoyang Tan, Hao Guan, William Smith, Shanmuganathan Raman, Mikhail Atroshenko, Xiaoyang Tan, Jonathan Weill, Shotaro Moriya, Alessandro Gentilini and Koki Saitoh who identified some of the problems below.

Simon Prince
June 10, 2015

Foreword, preface, acknowledgements

Page xvii Par. 4	'class of object' $\longrightarrow$ 'type of object'

Chapter 1

Chapter 2

Page 19 Sec. 3.2	'single parameter' $\longrightarrow$ 'single variable'

Chapter 3

Page 26 Prob. 3.4 — Not technically wrong, but it would have been more helpful if I'd given you the multivariate formulation of the exponential family:

$$Pr(x|\boldsymbol{\theta}) = a[\mathbf{x}]\exp[\mathbf{b}[\boldsymbol{\theta}]^T\mathbf{c}[\mathbf{x}] - d[\boldsymbol{\theta}]],$$

and asked you to find the functions $a[\mathbf{x}], \mathbf{b}[\boldsymbol{\theta}], \mathbf{c}[\mathbf{x}]$ and $d[\boldsymbol{\theta}]$.

Page 27 Prob. 3.7 — 'normal distribution' $\longrightarrow$ 'univariate normal distribution' and see errata to Problem 3.4.

Chapter 4

Page 39 Eqn 4.31 — Missing subscript k from the last λ on the third line of this equation

Page 40 Eqn 4.35 — This equation should not have tildes in either the numerator or denominator so that

$$\frac{N_k + \tilde{\alpha}_k}{\sum_{j=1}^{6}(N_j + \tilde{\alpha}_j)} \qquad \longrightarrow \qquad \frac{N_k + \alpha_k}{\sum_{j=1}^{6}(N_j + \alpha_j)}$$

Page 42 Prob. 4.2 — $\Gamma(\alpha) \longrightarrow \Gamma[\alpha]$

Chapter 5

Page 45 Parag. 1	'axes that are' $\longrightarrow$ 'axes are'
Page 51 Prob. 5.5	Round brackets starting after $\text{Norm}_{\mathbf{x}_1}$ in RHS of second equation should be square.
Page 51 Prob. 5.5	'Use Schur's complement.' $\longrightarrow$ 'Use the Schur complement identity.'

Chapter 6

Page 62 Discussion	$P(x|w) \longrightarrow Pr(x|w)$
Page 69 Eqn. 6.19	$\text{Norm}_{w_i}[\phi_0 + \phi_1 x] \longrightarrow \text{Norm}_{w_i}[\phi_0 + \phi_1 x, \sigma^2]$
Page 69 Prob. 6.7	'the locus of points where $Pr(w|x=0) = Pr(w|x=1)$' $\longrightarrow$ 'the locus of points where $Pr(w=0|x) = Pr(w=1|x)$'

Chapter 7

Page 79 Eqn 7.17 — The last term in the numerator and denominator of the top line should be conditioned on $\boldsymbol{\theta}^{[t]}$ and not joint. For example,

$$Pr(h_i = k, \boldsymbol{\theta}^{[t]}) \longrightarrow Pr(h_i = k|\boldsymbol{\theta}^{[t]}).$$

Page 90 Eqn. 7.34 — The covariance in the last line should be inverted so that

$$(\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi} + \mathbf{I}) \longrightarrow (\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi} + \mathbf{I})^{-1}$$

Page 92 Fig. 7.21 — Covariance in panel (b) is missing a term and should change as

$$(\boldsymbol{\phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\phi})^{-1} \longrightarrow (\boldsymbol{\phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\phi} + 1)^{-1}.$$

Similarly, the covariance in panel (d) should change as

$$(\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi})^{-1} \longrightarrow (\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi} + \mathbf{I})^{-1}$$

Page 93 Eqn 7.38 — Update for covariance is wrong and should be

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{I}\sum_{i=1}^{I} \text{diag}\left[(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T - \hat{\boldsymbol{\Phi}}\text{E}[\mathbf{h}_i](\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T\right],$$

Page 99 Eq. 7.52

$$\sum_{k=1}^{K} \lambda_{mk} \longrightarrow \sum_{k=1}^{K} \lambda_{km}$$

Page 105 Parag. 2 — ‘to the each example’ $\longrightarrow$ ‘to each example’

Page 106 Prob. 7.5 — $\text{Norm}_x[\mu, \sigma^2/\nu] \longrightarrow \text{Norm}_x[\mu, \sigma^2/h]$

Page 107 Prob. 7.8 — $E[(\mathbf{x}_i - E[\mathbf{x}_i])^2] \longrightarrow E[(\mathbf{x}_i - E[\mathbf{x}_i])(\mathbf{x}_i - E[\mathbf{x}_i])^T]$

Chapter 8

Page 110 Sec.8.1.2	'data $\mathbf{x}$ is' $\longrightarrow$ 'data $\mathbf{x}$ are'
Page117 Eq. 8.24	$\text{Norm}_w \longrightarrow \text{Norm}_{w^*}$
Page 122 Eq 8.34	Proportional sign $\propto$ in first line should be equals sign $=$
Page 125 Eq. 8.42	$\text{Norm}_{x_i} \longrightarrow \text{Norm}_{w_i}$
Page 125 Eq. 8.44	$\log[\sigma] \longrightarrow \log[\sigma^2]$
Page 125 Eq. 8.45	$\hat{\sigma}^2 = \frac{(\mathbf{w}-\mathbf{X}^T\mathbf{X}\boldsymbol{\psi})^T(\mathbf{w}-\mathbf{X}^T\mathbf{X}\boldsymbol{\psi})}{I}. \longrightarrow \hat{\sigma}^2 = \frac{(\mathbf{w}-\mathbf{X}^T\mathbf{X}\boldsymbol{\psi})^T(\mathbf{w}-\mathbf{X}^T\mathbf{X}\boldsymbol{\psi})}{I} = 0,$ where the variance will be zero when $D > I$.
Page 126 Parag. 1	The solution is $\longrightarrow$ When the matrix $\mathbf{X}$ is square and invertible, the solution is

Page 126 eq. 8.48 — The likelihood and evidence terms are wrong so that

$$\frac{Pr(\mathbf{X}|\mathbf{w},\boldsymbol{\psi},\sigma^2)Pr(\boldsymbol{\psi})}{Pr(\mathbf{X}|\mathbf{w},\sigma^2)} \longrightarrow \frac{Pr(\mathbf{w}|\mathbf{X},\boldsymbol{\psi},\sigma^2)Pr(\boldsymbol{\psi})}{Pr(\mathbf{w}|\mathbf{X},\sigma^2)}$$

Page 126 Eq. 8.51 — the term $\mathbf{w}^*$ in the second line of the equation should be just $\mathbf{w}$. So, the mean of the normal distribution changes as

$$\frac{1}{\sigma^2}\mathbf{x}^{*T}\mathbf{X}\mathbf{A}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{w}^* \longrightarrow \frac{1}{\sigma^2}\mathbf{x}^{*T}\mathbf{X}\mathbf{A}^{-1}\mathbf{X}^T\mathbf{X}\mathbf{w}$$

Page 127 Eq. 8.54 — The product term should be over the index i for examples rather than the index d for dimensions, so that

$$\prod_{d=1}^{D}\text{Gam}_{h_d}[\nu/2,\nu/2] \longrightarrow \prod_{i=1}^{I}\text{Gam}_{h_i}[\nu/2,\nu/2]$$

Page 128 Parag 2.	these may be also be $\longrightarrow$ these may also be
Page 129 Sec. 8.10.1	'circumvented this system' $\longrightarrow$ 'circumvented this problem'
Page 131 Prob. 8.2	'robust regression problem' $\longrightarrow$ 'robust regression model'
Page 132 Prob 8.8	All terms of form $q(h)$ should be replaced by $q(\phi)$.

Chapter 9

Page 141 Eq. 9.18	$\mathbf{x}^{*T}\boldsymbol{\Sigma}\mathbf{x} \longrightarrow \mathbf{x}^{*T}\boldsymbol{\Sigma}\mathbf{x}^{*}$
Page 144 Parag. 2	'data ... is high dimensional' $\longrightarrow$ 'data... are high dimensional'
Page152 Fig. 9.15	'e) After fitting ten RBFs. f) The data' $\longrightarrow$ 'f) After fitting ten RBFs. The data'
Page 156 Parag. 3	'the the posterior' $\longrightarrow$ 'the posterior'
Page 163 Sec 9.12.4	$w \in \{1,\ldots 3\} \longrightarrow w \in \{1,2,3\}$.
Page 168 Prob. 9.1	Missing label '(iii)'
Page 168 Prob. 9.3	'derivatives' $\longrightarrow$ 'derivative'

Chapter 10

Page 174 Fig 10.1	'proportions)and' $\longrightarrow$ 'proportions) and'
Page 185 Fig 10.10	'each marginal distributions' $\longrightarrow$ 'each marginal distribution'
Page 186 Eq. 10.17	Upper limit of product is wrong:

$$\prod_{n=1}^{I} \longrightarrow \prod_{n=1}^{N} \tag{1.1}$$

Chapter 11

Page 198 Parag. 2	'as a being' $\longrightarrow$ 'as being'

Chapter 12

Page 243 Fig 12.15	Missing unary cost in panel (b) in edge from vertex b_4 to b_5. Should be $U_b(4)$.
Page 253 Par. 2	'background ($w=0$) and background ($w=1$)' $\longrightarrow$ 'background ($w=0$) and foreground ($w=1$)'

Chapter 13

Page 284 Par. 3	128×8 vector $\longrightarrow$ 128×1 vector
Page 291 Par. 1	'represent the each' $\longrightarrow$ 'represent each'
Page 293 Prob. 13.2	$f[n,m] \longrightarrow f[m,n]$
Page 294 Prob. 13.9	'K-means algorithm' $\longrightarrow$ 'K-means algorithm with $K=2$ clusters'

Chapter 14

Page 311 Eq. 14.31	$\hat{\boldsymbol{\tau}} = \sum_{m=1}^{3}\sum_{n=1}^{3}\frac{\hat{\Omega}_{mn}}{\Omega_{mn}}\boldsymbol{\tau} \longrightarrow \hat{\boldsymbol{\tau}} = \sum_{m=1}^{3}\sum_{n=1}^{3}\frac{\hat{\Omega}_{mn}}{\Omega_{mn}}\boldsymbol{\tau}/9$
Page 316 Par. 4	'Figure 14.17' $\longrightarrow$ 'Figure 14.16'

Chapter 15

Page 329 Fig15.6	‘and shears and are’ $\longrightarrow$ ‘and shears are’

Chapter 16

Page 354 Par. 4	‘objective function is a based’ $\longrightarrow$ ‘objective function is based’
Page 369 Par. 3	‘with those in the first’ $\longrightarrow$ ‘with those in the second’
Page 381 Prob. 16.3	$(\mathbf{a}\times\mathbf{b}).\mathbf{c} \longrightarrow \mathbf{a}.(\mathbf{b}\times\mathbf{c}) = 0$

Chapter 17

Page 394 Eqn. 17.9	Missing close square bracket
Page 399 Eqn. 17.20	$\mathbf{w}_i = \boldsymbol{\mu} + \boldsymbol{\Phi}\mathbf{h}_i + \epsilon_i \longrightarrow \mathbf{w}_i = \boldsymbol{\mu} + \boldsymbol{\Phi}\mathbf{h}_i + \boldsymbol{\epsilon}_i$
Page 399 Par. 7	$\epsilon_i \longrightarrow \boldsymbol{\epsilon}_i$
Page 413 Eqn 17.43	$Pr(\mathbf{W}\vert\mathbf{X},\boldsymbol{\Phi},\sigma^2) \longrightarrow Pr(\mathbf{W}\vert\mathbf{H},\boldsymbol{\Phi},\sigma^2)$
Page 422 Prob. 17.7	Equation should read

$$\begin{aligned} \mathbf{W}\mathbf{W}^T &= \mathbf{U}\mathbf{L}^2\mathbf{U}^T \\ \mathbf{W}^T\mathbf{W} &= \mathbf{V}\mathbf{L}^2\mathbf{V}^T. \end{aligned}$$

Chapter 18

Page 425 Eqn. 18.4 — The covariance should be inverted so that

$$(\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}+\mathbf{I}) \longrightarrow (\boldsymbol{\Phi}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\Phi}+\mathbf{I})^{-1}$$

Page 425 Eqn 18.5 — Update for covariance is wrong and should be

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{I}\sum_{i=1}^{I} \mathrm{diag}\left[(\mathbf{x}_i-\hat{\boldsymbol{\mu}})(\mathbf{x}_i-\hat{\boldsymbol{\mu}})^T - \hat{\boldsymbol{\Phi}}\mathrm{E}[\mathbf{h}_i](\mathbf{x}_i-\hat{\boldsymbol{\mu}})^T\right],$$

Page 429 Fig. 18.6 — In caption: ‘c-l’ $\longrightarrow$ ‘c-j’

Page 429 Eqn 18.11 — Update for covariance is wrong and should be

$$\boldsymbol{\Sigma} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} \mathrm{diag}\left[(\mathbf{x}_{ij}-\hat{\boldsymbol{\mu}})(\mathbf{x}_{ij}-\hat{\boldsymbol{\mu}})^T - \hat{\boldsymbol{\Phi}}\mathrm{E}[\mathbf{h}_i](\mathbf{x}_{ij}-\hat{\boldsymbol{\mu}})^T\right],$$

Chapter 19

Page 481 Prob. 19.3 — The weights are missing from the equation which should become

$$Pr(\mathbf{w}_t|\mathbf{x}_{1\ldots t-1}) = \sum_{k=1}^{K}\lambda_k \text{Norm}_{\mathbf{w}_t}[\boldsymbol{\mu}_{+k}, \boldsymbol{\Sigma}_{+k}].$$

Page 481 Prob. 19.5 — 'max marginals solution' $\longrightarrow$ 'marginal posteriors'

Chapter 20

Page 492 Sec. 20.3.2 — 'can been used' $\longrightarrow$ 'can be used'

Page 493 Sec. 20.3.2 — 'in this mixtures' $\longrightarrow$ 'in these mixtures'

Page 495 Sec. 20.5 — 'class of model' $\longrightarrow$ 'class of models'

Page 498 Fig. 20.11 caption — 'equivalent of' $\longrightarrow$ 'equivalent to'

Page 499 Eq. 20.31 — The left hand side of the last equation should also be conditioned on w and becomes

$$Pr(\mathbf{x}_{ij}|p_{ij}=m, w_{ij}=n) = \text{Norm}_{\mathbf{x}_{ij}}[\boldsymbol{\mu}_n^{(w)} + \boldsymbol{\mu}_m^{(p)}, \boldsymbol{\Sigma}_n^{(w)} + \boldsymbol{\Sigma}_m^{(p)}].$$

Page 500 Sec. 20.7.1 — 'of the order of' $\longrightarrow$ 'on the order of'

Appendix A

Appendix B

Appendix C

Page 521 Sec. C.2.5 — 'class of square matrix' $\longrightarrow$ 'class of square matrices'

Page 521 Sec. C.2.5 — 'class of matrix' $\longrightarrow$ 'class of matrices'

Page 529 Sec. C.7.2 — $\text{argmax}_{\mathbf{b}}\,[\mathbf{Ab}] \longrightarrow \text{argmax}_{\mathbf{b}}\,[|\mathbf{Ab}|]$

Page 529 Sec. C.7.2 — $\text{argmin}_{\mathbf{b}}\,[\mathbf{Ab}] \longrightarrow \text{argmin}_{\mathbf{b}}\,[|\mathbf{Ab}|]$