2

Image Statistics

Exercise 1

From the definition of variance,

$$\begin{aligned}\mathrm{var}(G) &= \langle (G - \langle G\rangle)^2\rangle \\ &= \langle G^2 - 2G\langle G\rangle + \langle G\rangle^2\rangle \\ &= \langle G^2\rangle - 2\langle G\rangle^2 + \langle G\rangle^2 \\ &= \langle G^2\rangle - \langle G\rangle^2.\end{aligned}$$

Using this result,

$$\begin{aligned}\mathrm{var}(a_0 + a_1 G) &= \mathrm{var}(a_1 G) = \langle a_1^2 G^2\rangle - \langle a_1 G\rangle^2 \\ &= a_1^2(\langle G^2\rangle - \langle G\rangle^2) = a_1^2 \mathrm{var}(G).\end{aligned}$$

Exercise 2

Let $Y = |X|$ where X has the distribution $\Phi(x)$. For $y > 0$ we have

$$P(y) = \Pr(Y \le y) = \Pr(|X| \le y) = \Pr(-y \le X \le y) = \Phi(y) - \Phi(-y).$$

Hence

$$p(y) = \frac{d}{dy}P(y) = \phi(y) - \phi(-y) = 2\phi(y),$$

and clearly for $y \le 0$, $p(y) = 0$.

Exercise 3

Let X be standard normally distributed, i.e. with density function

$$p(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

We can't apply Theorem 2.1 directly to get the density function for X^2 because the function $u(x) = x^2$ is not monotonic wherever $p(x) \ne 0$. But consider the random variable

$$Y = Z^2, \quad \text{where} \quad Z = |X|.$$

We saw in the previous exercise that Z has the density function

$$f(z) = \begin{cases} 2\phi(z) & \text{for } z > 0 \\ 0 & \text{elsewhere.} \end{cases}$$

The function $y = u(z) = z^2$ is monotonic wherever $f(z) > 0$, so Theorem 2.1 holds. Inverting, $z = w(y) = y^{1/2}$. So for $y > 0$

$$g(y) = f(y^{1/2})\left|\frac{1}{2}y^{-1/2}\right| = \frac{2}{\sqrt{2\pi}}e^{-y/2}\frac{1}{2}y^{-1/2} = \frac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2},$$

which is the chi-square distribution, Equation (2.44), for $n = 1$.

Exercise 4

Given X_1 and X_2 independent and $\sim \mathcal{N}(0,1)$, we want the distribution of $Y = X_1 + X_2$. The joint density function for the random vector $\boldsymbol{X} = (X_1, X_2)^\top$ is, from independence,

$$f(x_1, x_2) = \frac{1}{\sqrt{2\pi}}e^{-x_1^2/2}\frac{1}{\sqrt{2\pi}}e^{-x_2^2/2} = \frac{1}{2\pi}e^{-x_1^2/2-x_2^2/2}$$

or

$$f(x_1, x_2) = \frac{1}{2\pi}e^{-(x_1+x_2)^2/2+x_1x_2}.$$

For fixed X_1, we can write $x_2 = w(y) = y - x_1$. Applying Theorem 2.1, Y has the density

$$g(y, x_1) = f(x_1, y - x_1)\left|\frac{\partial w}{\partial y}\right| = \frac{1}{2\pi}e^{-y^2/2+(y-x_1)x_1}.$$

The complete density for Y is then obtained by integrating over x_1. Dropping the subscript on x_1, this is

$$p(y) = \int_{-\infty}^{\infty} g(y, x)dx = \frac{1}{2\pi}e^{-y^2/2}\int_{-\infty}^{\infty} e^{(y-x)x}dx.$$

The integral is

$$\int_{-\infty}^{\infty} e^{(y-x)x}dx = \int_{-\infty}^{\infty} e^{xy-x^2+y^2/4-y^2/4}dx = e^{y^2/4}\int_{-\infty}^{\infty} e^{-(x-y/2)^2}dx.$$

Let $u = \sqrt{2}(x - y/2)$. Then the last expression above is

$$e^{y^2/4}\int_{-\infty}^{\infty} e^{-u^2/2}\frac{1}{\sqrt{2}}du = e^{y^2/4}\sqrt{\pi}.$$

Therefore

$$p(y) = \frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-y^2/4},$$

that is, $Y \sim \mathcal{N}(0, 2)$.

Exercise 5

We have, with Theorem 2.2,

$$\langle \bar{Z} \rangle = \left\langle \frac{1}{m} \sum_{i=1}^{m} Z_i \right\rangle = \frac{1}{m} m\mu = \mu,$$

and

$$\mathrm{var}(\bar{Z}) = \left(\frac{1}{m}\right)^2 \mathrm{var}\left(\sum_{i=1}^{m} Z_i\right) = \frac{1}{m^2} m\sigma^2 = \frac{1}{m}\sigma^2.$$

Exercise 6

The gamma function is

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

Integrating by parts ($\int v du = uv - \int u dv$) with

$$v = x^{\alpha-1}, \quad du = e^{-x} dx \Rightarrow u = -e^{-x},$$

we have

$$\begin{aligned} \Gamma(\alpha) &= -e^{-x} x^{\alpha-1} \Big|_0^\infty + \int_0^\infty e^{-x} (\alpha-1) x^{\alpha-2} dx \\ &= (\alpha-1) \int_0^\infty x^{\alpha-2} e^{-x} = (\alpha-1)\Gamma(\alpha-1). \end{aligned}$$

Exercise 7

(a) From Equation (2.33)

$$\langle X \rangle = \int_0^\infty x \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx.$$

Let $y = x/\beta$. Then

$$\begin{aligned} \langle X \rangle &= \int_0^\infty \beta y \frac{1}{\beta^\alpha \Gamma(\alpha)} \beta^{\alpha-1} y^{\alpha-1} e^{-y} \beta dy \\ &= \beta \frac{1}{\Gamma(\alpha)} \int_0^\infty y^\alpha e^{-y} dy = \beta \frac{1}{\Gamma(\alpha)} \Gamma(\alpha+1) = \alpha\beta. \end{aligned}$$

Similarly

$$\begin{aligned} \langle X^2 \rangle &= \int_0^\infty (\beta y)^2 \frac{1}{\beta^\alpha \Gamma(\alpha)} \beta^{\alpha-1} y^{\alpha-1} e^{-y} \beta dy \\ &= \beta^2 \frac{1}{\Gamma(\alpha)} \int_0^\infty y^{\alpha+1} e^{-y} dy \\ &= \beta^2 \frac{1}{\Gamma(\alpha)} \Gamma(\alpha+2) = \beta^2 \frac{1}{\Gamma(\alpha)} (\alpha+1)\Gamma(\alpha+1) = \alpha(\alpha+1)\beta^2. \end{aligned}$$

Therefore

$$\mathrm{var}(X^2) = \langle X^2\rangle - \langle X\rangle^2 = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2.$$

(b) For exponentially distributed Z_1 and Z_2, let $Z = Z_1 + Z_2$. Then the distribution function for Z is

$$\begin{aligned}
P(z) = \Pr(Z_1 + Z_2 < z) &= \int_0^z \int_0^{z-z_2} \frac{1}{\beta^2} e^{z_1/\beta} e^{z_2/\beta} dz_1 dz_2 \\
&= \frac{1}{\beta^2} \int_0^z e^{z_2/\beta} \int_0^{z-z_2} e^{z_1/\beta} dz_1 \; dz_2 \\
&= \frac{1}{\beta^2} \int_0^z e^{z_2/\beta} \left[-\beta e^{z_1/\beta} \Big|_0^{z-z_2} \right] dz_2 \\
&= \frac{1}{\beta^2} \int_0^z e^{z_2/\beta} \left[-\beta e^{-z/\beta} e^{z_2/\beta} + \beta \right] dz_2 \\
&= \frac{1}{\beta^2} \int_0^z \left[-\beta e^{-z/\beta} + \beta e^{-z_2/\beta} \right] dz_2 \\
&= \frac{1}{\beta^2} \left[-\beta z e^{-z/\beta} + \beta \left[-\beta e^{z_2/\beta} \right]_0^z \right] \\
&= \frac{1}{\beta^2} \left[-\beta z e^{-z/\beta} + \beta \left[-\beta (e^{-z/\beta} - 1) \right] \right] \\
&= -\frac{1}{\beta} z e^{-z/\beta} - e^{-z/\beta} + 1.
\end{aligned}$$

Differentiating,

$$\begin{aligned}
p(z) = \frac{d}{dz} P(z) &= -\frac{1}{\beta} \left[e^{-z/\beta} - \frac{z}{\beta} e^{-z/\beta} \right] + \frac{1}{\beta} e^{-z/\beta} \\
&= \frac{z}{\beta^2} e^{-z/\beta} = \frac{1}{\beta^2 \Gamma(2)} z e^{-z/\beta}
\end{aligned}$$

since $\Gamma(2) = 1$.

Exercise 8

For two dimensions, the covariance is

$$\begin{aligned}
&\mathrm{cov}(\boldsymbol{a}^\top \boldsymbol{G}, \boldsymbol{b}^\top \boldsymbol{G}) = \mathrm{cov}(a_1 G_1 + a_2 G_2, b_1 G_1 + b_2 G_2) \\
&\quad = a_1 b_1 \mathrm{var}(G_1) + a_1 b_2 \mathrm{cov}(G_1 G_2) + a_2 b_1 \mathrm{cov}(G_2, G_1) + a_2 b_2 \mathrm{var}(G_2),
\end{aligned}$$

which by inspection is the same as

$$(a_1, a_2) \begin{pmatrix} \mathrm{var}(G_1) & \mathrm{cov}(G_1, G_2) \\ \mathrm{cov}(G_2, G_1) & \mathrm{var}(G_2) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{b}.$$

More generally,

$$\mathrm{cov}(\boldsymbol{a}^\top \boldsymbol{G}, \boldsymbol{b}^\top \boldsymbol{G}) = \langle \boldsymbol{a}^\top \boldsymbol{G} (\boldsymbol{b}^\top \boldsymbol{G})^\top \rangle = \boldsymbol{a}^\top \langle \boldsymbol{G}\boldsymbol{G}^\top \rangle \boldsymbol{b} = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{b}.$$

Exercise 9

For $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}$, $|\boldsymbol{\Sigma}| = (\sigma^2)^N$, and the multivariate density simplifies as follows

$$\begin{aligned} p(\boldsymbol{g}) &= \frac{1}{(2\pi)^{N/2}\sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\boldsymbol{g}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{g}-\boldsymbol{\mu})\right) \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^N} \exp\left(-\|\boldsymbol{g}-\boldsymbol{\mu}\|^2/2\sigma^2\right) \\ &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-(g_i-\mu_i)^2/2\sigma^2\right). \end{aligned}$$

For $N = 1$, therefore,

$$p(g) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-(g-\mu)^2/2\sigma^2\right).$$

The mean of G is given by

$$\langle G \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} g \cdot e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} dg$$

which we can write in the form

$$\langle G \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{g-\mu}{\sigma} e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} dg + \frac{\mu}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} dg.$$

The second term is just $\mu \int_{-\infty}^{\infty} p(g) dg = \mu$. The first term vanishes. That is, making the change of variable substitution

$$y = \frac{1}{2}\left(\frac{g-\mu}{\sigma}\right),$$

we have

$$\int_{\mu}^{\infty} \frac{g-\mu}{\sigma} e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} dg = \sigma \int_0^{\infty} e^{-y} dy = \sigma$$

and similarly

$$\int_{-\infty}^{\mu} \frac{g-\mu}{\sigma} e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} dg = -\sigma.$$

Thus $\langle G \rangle = \mu$.

Exercise 10

The covariance matrix estimate can be written as

$$\boldsymbol{s} = \frac{m}{m-1} \cdot \frac{1}{m} \sum_{i=1}^m (\boldsymbol{z}_i - \bar{\boldsymbol{z}})(\boldsymbol{z}_i - \bar{\boldsymbol{z}})^\top.$$

Expanding,

$$\begin{aligned}\boldsymbol{s} &= \frac{m}{m-1}\left[\frac{1}{m}\sum_i \boldsymbol{z}_i\boldsymbol{z}_1^\top - 2\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^\top + \frac{1}{m}\cdot m\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^\top\right]\\ &= \frac{m}{m-1}\left[\frac{1}{m}\boldsymbol{\mathcal{Z}}^\top\boldsymbol{\mathcal{Z}} - \bar{\boldsymbol{z}}\bar{\boldsymbol{z}}^\top\right]\end{aligned}$$

With Equation (2.52) we can write

$$\begin{aligned}\boldsymbol{s} &= \frac{m}{m-1}\left[\frac{1}{m}\boldsymbol{\mathcal{Z}}^\top\boldsymbol{\mathcal{Z}} - \frac{1}{m^2}\boldsymbol{\mathcal{Z}}^\top\mathbf{1}_m\mathbf{1}_m^\top\boldsymbol{\mathcal{Z}}\right]\\ &= \frac{1}{m-1}\left[\boldsymbol{\mathcal{Z}}^\top\boldsymbol{\mathcal{Z}} - \frac{1}{m}\boldsymbol{\mathcal{Z}}^\top\mathbf{1}_m\mathbf{1}_m^\top\boldsymbol{\mathcal{Z}}\right]\\ &= \frac{1}{m-1}\boldsymbol{\mathcal{Z}}^\top\left[\boldsymbol{I}_{mm} - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top\right]\boldsymbol{\mathcal{Z}} = \frac{1}{m-1}\boldsymbol{\mathcal{Z}}\boldsymbol{H}\boldsymbol{\mathcal{Z}}.\end{aligned}$$

Clearly, $\boldsymbol{H}^\top = \boldsymbol{H}$, so $\boldsymbol{H}$ is symmetric. Also

$$\begin{aligned}\boldsymbol{H}\boldsymbol{H} &= \left[\boldsymbol{I}_{mm} - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top\right]\left[\boldsymbol{I}_{mm} - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top\right]\\ &= \boldsymbol{I}_{mm} - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top + \frac{1}{m^2}\mathbf{1}_m\mathbf{1}_m^\top\mathbf{1}_m\mathbf{1}_m^\top\\ &= \boldsymbol{I}_{mm} - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^\top = \boldsymbol{H},\end{aligned}$$

since $\mathbf{1}_m^\top\mathbf{1}_m = m$. Let $\boldsymbol{x}$ be any m-component vector. Then with the above result

$$\boldsymbol{x}^\top\boldsymbol{s}\boldsymbol{x} = \frac{1}{m-1}\boldsymbol{x}^\top\boldsymbol{\mathcal{Z}}^\top\boldsymbol{H}\boldsymbol{H}\boldsymbol{\mathcal{Z}}\boldsymbol{x} = \frac{1}{m-1}\boldsymbol{y}^\top\boldsymbol{y} \ge 0,$$

where $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{\mathcal{Z}}\boldsymbol{x}$. Therefore $\boldsymbol{s}$ is positive semi-definite.

Exercise 11

With $s_i = \sqrt{s_{ii}}$ Equation (2.54) is

$$\begin{aligned}\boldsymbol{d}^{-1/2}\boldsymbol{s}\boldsymbol{d}^{-1/2} &= \begin{pmatrix}1/s_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1/s_m\end{pmatrix}\begin{pmatrix}s_{11} & \cdots & s_{1m}\\ \vdots & \ddots & \vdots\\ s_{m1} & \cdots & s_{mm}\end{pmatrix}\begin{pmatrix}1/s_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1/s_m\end{pmatrix}\\ &= \begin{pmatrix}s_{11}/s_1^2 & \cdots & s_{1m}/s_1s_m\\ \vdots & \ddots & \vdots\\ s_{m1}/s_ms_1 & \cdots & s_{mm}/s_m^2\end{pmatrix} = \boldsymbol{r}.\end{aligned}$$

Exercise 12

Let A_i represent the situation "auto is behind door i". The *a priori* probabilities are $\Pr(A_i) = 1/3$, $i = 1, 2, 3$. Let O_i be the observation "quizmaster opens door i". Suppose the contestant chooses door 2 and the quizmaster opens door 1. Then we have the following conditional probabilities:

$$\begin{aligned} \Pr(O_1 \mid A_1) &= 0 \quad \text{quizmaster won't give the car away.} \\ \Pr(O_1 \mid A_2) &= 1/2 \quad \text{quizmaster is indifferent.} \\ \Pr(O_1 \mid A_3] &= 1 \quad \text{quizmaster has no choice.} \end{aligned}$$

Now apply Bayes' Theorem to find the *a posteriori* probability that the auto is behind door 2, given the observation:

$$\begin{aligned} \Pr(A_2 \mid O_1) &= \frac{\Pr(O_1 \mid A_2)\Pr(A_2)}{\Pr(O_1 \mid A_2)\Pr(A_2) + \Pr(O_1 \mid A_3)\Pr(A_3)} \\ &= \frac{(1/2)(1/3)}{(1/2)(1/3) + (1)(1/3)} = 1/3, \end{aligned}$$

whereas, with the same argument,

$$\Pr(A_3 \mid O_1) = 2/3.$$

The contestant would therefore be well-advised to switch to door 3.

Exercise 13

```
X = randomu(seed,1000,/normal)
Y = randomu(seed,1000,/normal)
tm = tm_test(X,Y)
PRINT, 'p-value␣=␣', tm[1]
fv = fv_test(X,Y)
PRINT, 'p-value␣=␣', fv[1]

p-value =      0.246175
p-value =      0.477225
```

Exercise 14

$$\begin{aligned} \frac{L(\mu_0)}{L(\hat{\mu})} &= \exp\left(-\sum_\nu (z(\nu) - \mu_0)^2/2\sigma^2\right) \Big/ \exp\left(-\sum_\nu (z(\nu) - \bar{z})^2/2\sigma^2\right) \\ &= \exp\left(-\frac{1}{2\sigma^2}\sum_\nu (z(\nu) - \mu_0)^2 + (z(\nu) - \bar{z})^2\right) \\ &= \exp\left(-\frac{1}{2\sigma^2/m}(\bar{z} - \mu_0)^2\right) \le k \end{aligned}$$

Exercise 15

Differentiating Equation (2.84):

$$0 = \frac{\partial z}{\partial a} = -\frac{2}{\sigma^2}\sum_i (y_i - a - bx_i)$$

$$0 = \frac{\partial z}{\partial b} = -\frac{2}{\sigma^2}\sum_i x_i(y_i - a - bx_i).$$

From the first equation we get immediately

$$a = \bar{y} - b\bar{x},$$

where $\bar{x}$ and $\bar{y}$ are given by Equation (2.86). Substituting for a in the second equation, we have

$$0 = \sum_i x_i y_i - (\bar{y} - b\bar{x})\sum_i x_i - b\sum_i x_i^2.$$

Re-arranging:

$$b\left(\sum_i x_i^2 - \bar{x}\sum_i x_i\right) = \sum_i x_i y_i - \bar{y}\sum_i x_i.$$

The expression in brackets on the left is just

$$\sum_i x_i^2 - m\bar{x}^2 = \sum_i (x_i - \bar{x})^2 = m s_{xx},$$

as can easily be seen by expansion. Similarly the right hand side is

$$\sum_i (x_i - \bar{x})(y_i - \bar{y}) = m s_{xy}.$$

Hence

$$b = \frac{s_{xy}}{s_{xx}}.$$

To show that these values minimize Equation (2.84), we require the Hessian matrix:

$$\frac{\partial^2 z}{\partial a^2} = \frac{2m}{\sigma^2}$$

$$\frac{\partial^2 z}{\partial b^2} = \frac{2}{\sigma^2}\sum_i x_i^2$$

$$\frac{\partial^2 z}{\partial b \partial a} = \frac{\partial^2 z}{\partial a \partial b} = \frac{2}{\sigma^2}\sum_i x_i$$

Listing 2.1: Solution to Exercise 18.

```
PRO solution2_18

envi_select, title='Choose␣multispectral␣image', $
             fid=fid, dims=dims,pos=pos
IF (fid EQ -1) THEN BEGIN
   PRINT, 'cancelled'
   RETURN
ENDIF
envi_file_query, fid, fname=fname
num_cols = dims[2]-dims[1]+1
num_rows = dims[4]-dims[3]+1
num_pixels = num_cols*num_rows
image = fltarr(2,num_pixels)
FOR i=0,1 DO image[i,*] = $
   envi_get_data(fid=fid,dims=dims,pos=pos[i])
X = image[0,*]
Y = image[1,*]
result = regress(X[*],Y[*],const=a)
b = result[0]
; plot results to postcript file
thisDevice =!D.Name
set_plot, 'PS'
Device,Filename='c:\temp\fig_sol_2.eps',xsize=6,ysize=4,$
   /inches,/encapsulated
PLOT,image[0,*],image[1,*],psym=3
oplot, [min(X),max(X)], [a+b*min(X),a+b*max(X)]
device,/close_file
set_plot, thisDevice

END
```

The Hessian matrix is thus

$$\boldsymbol{H} = \frac{2m}{\sigma^2}\begin{pmatrix} 1 & \bar{x} \\ \bar{x} & s_{xx} + \bar{x}^2 \end{pmatrix}.$$

For any vector $\boldsymbol{y} \neq 0$ we then have

$$\boldsymbol{y}^\top \boldsymbol{H}\boldsymbol{y} = (y_1 + \bar{x}y_2)^2 + s_{xx}y_2^2 > 0,$$

so $\boldsymbol{H}$ is positive definite and (a, b) is a minimum of Equation (2.84).

Exercise 16

We require the derivative of a with respect to y_i, where

$$a = \bar{y} - b\bar{x} = \frac{1}{m}\sum_i y_i - \frac{s_{xy}}{s_{xx}}\bar{x}.$$

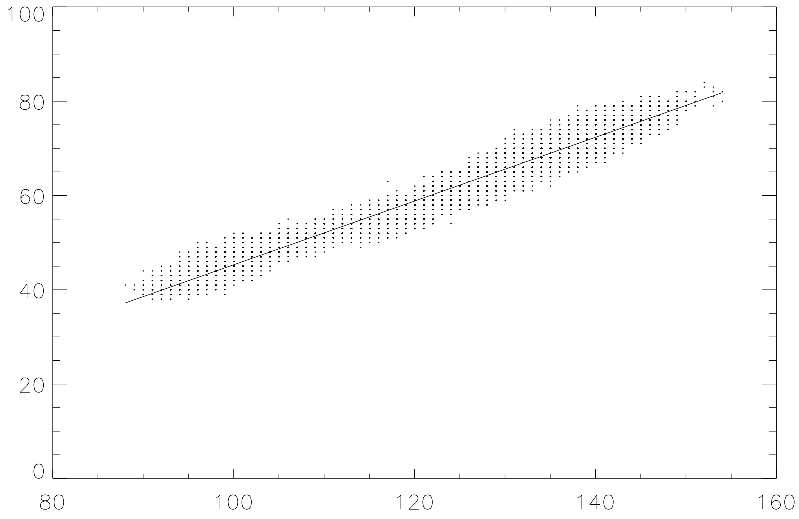

FIGURE 2.1
Regression of band 2 on band 1.

We have

$$\frac{\partial a}{\partial y_i} = \frac{1}{m} - \frac{\bar{x}}{s_{xx}} \frac{\partial s_{xy}}{\partial y_i}$$

and

$$\begin{aligned} \frac{\partial s_{xy}}{\partial y_i} &= \frac{1}{m} \left(x_i - \bar{x} - \sum_j (x_j - \bar{x}) \frac{\partial \bar{y}}{\partial y_i} \right) \\ &= \frac{1}{m}(x_i - \bar{x}). \end{aligned}$$

Hence

$$\frac{\partial a}{\partial y_i} = \frac{1}{m} \left(1 - \frac{\bar{x}}{s_{xx}} (x_i - \bar{x}) \right)$$

and

$$\left(\frac{\partial a}{\partial y_i} \right)^2 = \frac{1}{m^2} \left(1 - \frac{2\bar{x}}{s_{xx}} (x_i - \bar{x}) + \frac{\bar{x}^2}{s_{xx}^2} (x_i - \bar{x})^2 \right).$$

Summing over i gives

$$\begin{aligned} \sum_i \left(\frac{\partial a}{\partial y_i} \right)^2 &= \frac{1}{m^2} \left(m + 0 + \frac{\bar{x}^2}{s_{xx}^2} m s_{xx} \right) \\ &= \frac{1}{m} \left(\frac{s_{xx} + \bar{x}^2}{s_{xx}} \right) = \frac{\sum_i x_i^2}{m^2 s_{xx}}, \end{aligned}$$

from which the expression for σ_a^2 in Equation (2.87) readily follows.

Exercise 17

From Equation (2.94),

$$z(\boldsymbol{w}) = \frac{1}{\sigma^2}(\boldsymbol{y}^\top \boldsymbol{y} - \boldsymbol{y}^\top \boldsymbol{\mathcal{X}} \boldsymbol{w} - \boldsymbol{w}^\top \boldsymbol{\mathcal{X}} \boldsymbol{y} + \boldsymbol{w}^\top \boldsymbol{\mathcal{X}}^\top \boldsymbol{\mathcal{X}} \boldsymbol{w}),$$

so that

$$\frac{\partial z(\boldsymbol{w})}{\partial \boldsymbol{w}} = 0 = \frac{1}{\sigma^2}(-\boldsymbol{\mathcal{X}}^\top \boldsymbol{y} - \boldsymbol{\mathcal{X}}^\top \boldsymbol{y} + 2\boldsymbol{\mathcal{X}}^\top \boldsymbol{\mathcal{X}} \boldsymbol{w})$$

from which Equation (2.95) follows.

Exercise 18

An IDL routine for the regression is shown in Listing 2.1, see also Figure 2.1.

Exercise 19

The numerator of right hand side of Equation (2.90) becomes

$$\begin{aligned}
&\sum_\nu (y(\nu) + \hat{y}(\nu) - \hat{y}(\nu) - \bar{y})(\hat{y}(\nu) - \bar{y}) \\
&= \sum_\nu (y(\nu) - \hat{y}(\nu))(\hat{y}(\nu) - \bar{y}) + \sum_\nu (\hat{y}(\nu) - \bar{y})^2 \\
&= \sum_\nu (\hat{y}(\nu) - \bar{y})^2,
\end{aligned}$$

since the term $\sum_\nu (y(\nu) - \hat{y}(\nu))(\hat{y}(\nu) - \bar{y})$ vanishes, see Section 2.6.2. Including the denominator then gives the desired result.

Exercise 20

From Equations (2.103) and (2.104),

$$\hat{\boldsymbol{w}} = \frac{1}{\lambda}(\boldsymbol{\mathcal{X}}^\top \boldsymbol{y} - \boldsymbol{\mathcal{X}}^\top \boldsymbol{\mathcal{X}} \boldsymbol{w})$$

Solving for $\hat{\boldsymbol{w}}$ gives Equation (2.105).

Exercise 21

We want to show that

$$-\int p(x) \ln \left[\frac{q(x)}{p(x)}\right] dx \geq 0$$

when $q(x) \neq p(x)$. From Jensen's inequality, Equation (2.117), identify $f(x)$ with the convex function $-\ln(x)$ and $g(x)$ with the ratio $q(x)/p(x)$ to get

$$-\int p(x) \ln \left[\frac{q(x)}{p(x)} \right] dx \geq -\ln \left(\int q(x) dx \right).$$

But, since $q(x)$ is also a probability density function, the integral on the right hand side is one, and its logarithm vanishes.