STAT 8230 - Homework #2 - Solution

Fall 2011

#1. See the file hwk2-1.pdf on the course web site. This file contains a print-out of an S-PLUS script, its output and the associated graphs. Refer to these results.

a) In hwk2-1.ssc I first plotted log(Volume) $\equiv$ logV versus both logH (log height) and logD (log diameter). These plots show positive associations that look roughly linear. I fit model (*) using the lm( ) function and stored the fitted model in m1. A summary of the fitted model appears on p. 2 of hwk2-1.pdf. The parameter estimates and standard errors from this model are as follows

	parameter	estimate	s.e.
(intercept)	β_1	-6.63	0.800
(coef on logD)	β_2	1.98	0.075
(coef on logH)	β_3	1.12	0.204
(error SD)	σ	0.0814	

$\hat{\beta}_1 = -1.70\,(.882)$ if diameter is transformed to feet first

The overall regression F test is highly significant ($F_{2,28} = 613.2$, $p < .0001$) and each of the individual predictors (logD, logH) have a significant positive effect on logV ($p < .0001$ in each case). The R^2 from the fitted model is $R^2 = .978$ indicating that the model explains 97.8% of the variability in logV.

b.) Model (**) was fit as M2 using the nls() function. Starting values were taken to be

$$\hat{\underline{\theta}}^{\circ} = \begin{pmatrix} \exp(\hat{\beta}_1) \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}.$$

The fitted model is summarized on the bottom of p. 2 of hwk2-1.pdf.

Parameter	Estimate	S.E.
θ_1	.00145	.00137
θ_2	2.00	.0821
θ_3	1.08	.242
σ	2.53	

c) The 3rd + 4th plots on p. 3 of hwk2-1.pdf show the Pearson residuals vs fitted values for models M1 + M2. The residuals from model M2 appear to have increasing spread as the fitted values increase (a $<$ shape).

Therefore I would say that the homoscedasticity assumption apears to be more appropriate for model (*) than model (**).

(This part is not necessary to include in your answer to the problem.)

\# 1 (continued) If ~~trees~~ cherry tree boles are indeed shaped like cones, then we would expect

$\beta_1 \approx \log\left(\pi/12^3\right)$, $\beta_2 \approx 2$, $\beta_3 \approx 1$. Therefore,

this is $\pi/(12)^3$ rather than $\pi/12$ because diameter was in inches. We need to transform to feet (replace d by $d/12$)

in this problem it is of interest to test the hypothesis

$$H_0: \underline{\beta} = \begin{pmatrix} \log(\pi/12^3) \\ 2 \\ 1 \end{pmatrix} \quad \text{vs} \quad H_1: \underline{\beta} \neq \begin{pmatrix} \log \pi/12^3 \\ 2 \\ 1 \end{pmatrix}$$

This can be done using the F test given on p. 23 of our class notes:

$$F = \frac{(\hat{\beta} - \beta_0)^T (X^T X)(\hat{\beta} - \beta_0)}{p s^2} \sim F(p, n-p) \text{ under } H_0: \underline{\beta} = \beta_0$$

In this problem, we obtain

$F = 31.05$ which has a p-value $< .0001$ when compared w/ its null distribution

$$F(3, 28)$$

~~So, while $\hat{\beta}_2$, $\hat{\beta}_3$ are close the the null values 2 & 1, respectively, $\hat{\beta}_1$ is not.~~ This leads to rejection of H_0. The volume formula for a cone is not an adequate description of the mean volume of cherry trees.

#2 (2.3 in Bates & Watts)

See the file hwk2-2.pdf on the course web site. This file contains an S-PLUS script hwk2-2.ssc, its output, and the associated graphs.

a+b) The design $\underline{x} = (7, 28)^T$ has less intrinsic and less parameter-effects nonlinearity than the design $\underline{x} = (4, 41)^T$. This is so because the former design has an expectation surface that is less curved and more evenly-spaced with respect to θ-values. The design $\underline{x} = (12, 14)^T$ has less nonlinearity of both types for the same reasons.

c.) A design with $\underline{x} = (c, c)^T$ for any constant c would have zero intrinsic nonlinearity because the expectation ~~surface~~ function would have the form

$$60 + 70\left(e^{-c}\right)^{\theta} = 60 + 70\,\phi \quad \text{where } \phi = e^{-c\theta},$$

which is linear in ϕ.

d.) Such a design would not have zero parameter-effects nonlinearity because it does under the ϕ parameterization and unit changes in ϕ do not correspond to unit changes in θ. A picture of the (intrinsically linear) expectation surface for $c = 13$ appears in hwk2-2.pdf.

3

(2.5) a) See hwk 2-3.pdf ~~and its output hwk 2-3.out.~~
~~All of the results can be found in hwk 2-3.out.~~

The exp'tion vector is $(.0537, .8921, \ldots, 2.0740)^T$, the residual vector is $(-.03369, \ldots, .09599)^T$,

$$S(\underline{\theta}^0) = .07472, \quad \underline{V}^0 = \begin{pmatrix} .2289 & 1.6963 \\ \vdots & \\ .9427 & 1.3859 \end{pmatrix},$$

b.) $$\underline{\delta}^0 = \begin{pmatrix} .2570 \\ -.06501 \end{pmatrix}, \quad \underline{\theta}^1 = \begin{pmatrix} 2.2642 \\ .2437 \end{pmatrix}, \quad S(\underline{\theta}^1) = .05402, \text{ for } \lambda = \tfrac{1}{4}$$

$$\underline{\theta}^1 = \begin{pmatrix} 2.3285 \\ .2275 \end{pmatrix}, \quad S(\underline{\theta}^1) = .04137, \text{ for } \lambda = \tfrac{1}{2}$$

$$\underline{\theta}^1 = \begin{pmatrix} 2.4570 \\ .1950 \end{pmatrix}, \quad S(\underline{\theta}^1) = .04865, \text{ for } \lambda = 1$$

Since $S(\underline{\theta}^1) < S(\underline{\theta}^0)$ when $\lambda = 1$ its not necessary to use a fractional step size.

The option trace=T in nls() yields details of the G-N iterations that confirm the above results.

~~See hwk2-4.q, hwk2-1.out, hwk2-4.pdf~~

Extra problem:

Step 0 $\qquad \underline{\hat{\Theta}}^0 = .25, \quad \underline{\eta}(\underline{\hat{\Theta}}^0) = \begin{pmatrix} 1 - e^{-\hat{\Theta}^0} \\ 1 - e^{-5\hat{\Theta}^0} \end{pmatrix} = \begin{pmatrix} 1 - e^{-.25} \\ 1 - e^{-1.25} \end{pmatrix} = \begin{pmatrix} .2212 \\ .7135 \end{pmatrix}$

$\underline{z}^0 = \underline{y} - \underline{\eta}(\underline{\hat{\Theta}}^0) = \begin{pmatrix} .5 - .2212 \\ .5 - .7135 \end{pmatrix} = \begin{pmatrix} .2788 \\ -.2135 \end{pmatrix}$

$\boxed{\text{See hwk2-4.R and its output, hwk2-4.pdf}}$

$$\frac{\partial \eta(\theta)}{\partial \theta} = x e^{-\theta x} \Rightarrow V^0 = \begin{pmatrix} 1\, e^{-.25} \\ 5\, e^{-1.25} \end{pmatrix} = \begin{pmatrix} .7788 \\ 1.4325 \end{pmatrix}$$

$$\delta^0 = \{(V^0)^T V^0\}^{-1} V^0 z_0 = -.03337$$

The point on the tangent plane at $\hat{\theta}^0 = .25$ closest to $y = (.5, .5)^T$ is $\eta(\hat{\theta}^0) + V^0 \delta^0$

$$= \begin{pmatrix} .2212 \\ .7135 \end{pmatrix} + -.03337 \begin{pmatrix} .7788 \\ 1.4325 \end{pmatrix} = \begin{pmatrix} .1952 \\ .6657 \end{pmatrix}$$ (= tanO in hwk2-4.9 marked with a * on plots)

$$\hat{\theta}^1 = \hat{\theta}^0 + \delta^0 = .25 + (-.03337) = .2166$$

$$\eta(\hat{\theta}^1) = \begin{pmatrix} 1 - e^{-.2166} \\ 1 - e^{-.2166(5)} \end{pmatrix} = \begin{pmatrix} .1948 \\ .6615 \end{pmatrix}$$ (marked w/ a Δ on plots)

$$z^1 = y - \eta(\hat{\theta}^1) = \begin{pmatrix} .5 - .1948 \\ .5 - .6615 \end{pmatrix} = \begin{pmatrix} .3052 \\ -.1615 \end{pmatrix}$$

$$V^1 = \begin{pmatrix} 1\, e^{-.2166} \\ 5\, e^{-.2166(5)} \end{pmatrix} = \begin{pmatrix} .8052 \\ 1.6926 \end{pmatrix}$$

$$\delta^1 = \{(V^1)^T V^1\}^{-1} (V^1)^T z^1 = -.007841$$

The point on the tangent plane at $\hat{\theta}^1 = .2166$ closest to y is $\eta(\hat{\theta}^1) + V^1 \delta^1 = \begin{pmatrix} .1948 \\ .6615 \end{pmatrix} + -.007841 \begin{pmatrix} .8052 \\ 1.6926 \end{pmatrix} = \begin{pmatrix} .8052 \\ 1.6926 \end{pmatrix}$

(shown on 2nd plot as an *)

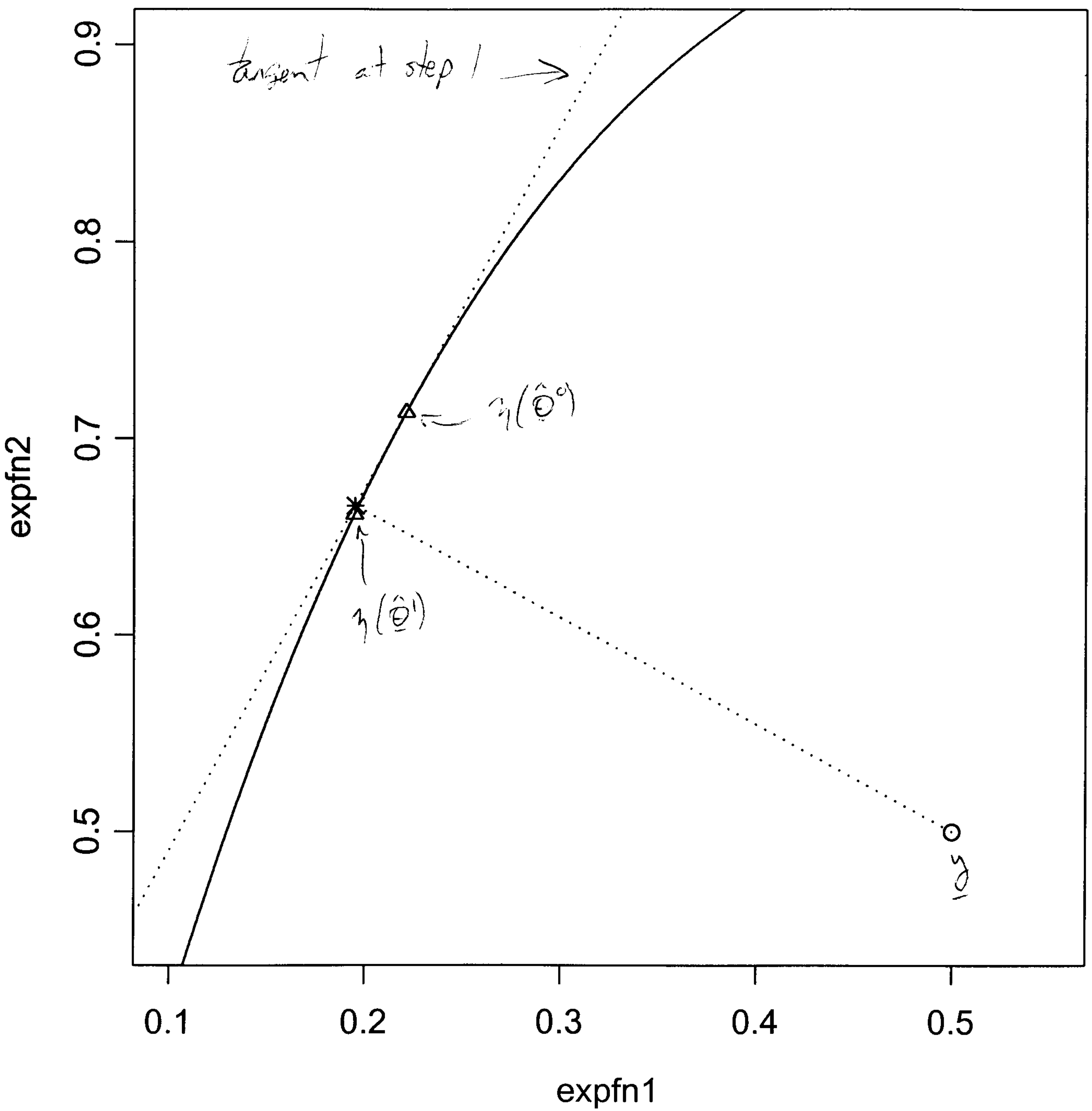
$$\hat{\Theta}^2 = \hat{\Theta}^1 + \delta^1 = .2166 - .007841 = .2088$$

$$z(\hat{\Theta}^2) = \begin{pmatrix} 1 - e^{-.2088} \\ 1 - e^{-.2088(5)} \end{pmatrix} = \begin{pmatrix} .8116 \\ 1.7603 \end{pmatrix}$$ shown on second plot as a Δ

These iterations are confirmed w/ the trace=T option in nls()

The following plot (next page) shows the 1st iteration.
The second plot shows the 1st two iterations in close-up! ~~These plots are attached as well as available electronically on the course web site.~~

Step 1, G-N Algorithm



Steps 1-2, G-N Algorithm