Topics in Statistical Sciences 2 – Exam exercise 4

Søren Højsgaard and Torben Tvedebrink

Version: 01/12/2017 11:16

This exercise is about the *Nonlinear Regression* as discussed in lectures 10 – 12 of Topics in Statistical Sciences 2. During the oral exam you will have 20 min to present the exercise. You decide what topics to cover and how to present them, however, we will ask questions to any part of the covered curricula, exercise and presentation.

1. In the first lecture we discussed the convergence criterion, *relative offset convergence criterion*:

$$\frac{\left\| {\mathbf{Q}_1^{(i)}}^\top \{\boldsymbol{y} - \boldsymbol{\eta}(\beta^{(i)})\} \right\|_2 / \sqrt{p}}{\left\| {\mathbf{Q}_2^{(i)}}^\top \{\boldsymbol{y} - \boldsymbol{\eta}(\beta^{(i)})\} \right\|_2 / \sqrt{n-p}} \tag{1}$$

where $\mathbf{Q}^{(i)} = \begin{bmatrix} \mathbf{Q}_1^{(i)} & \mathbf{Q}_2^{(i)} \end{bmatrix}$ from the QR-factorisation of $\mathbf{V}^{(i)}$ (*NB!* More superscripts than on slides).

This is related to the geometry of linear models, $\boldsymbol{y} = \mathbf{X}\beta + \boldsymbol{\varepsilon}$. For the QR-factorisation where $\mathbf{X} = \mathbf{QR}$, $\mathbf{Q}$ is a orthogonal $n \times n$ with $\mathbf{Q} = [\mathbf{Q}_1\ \mathbf{Q}_2]$, where two parts $\mathbf{Q}_1$ and $\mathbf{Q}_2$ is $n \times p$ and $n \times (n-p)$, respectively. The $\mathbf{R}$ matrix is $n \times p$, and has an upper triangular matrix R_1 in the first p rows, and zeros below, $\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \boldsymbol{O} \end{bmatrix}$. Hence, $\mathbf{X} = \mathbf{QR} = \mathbf{Q}_1\mathbf{R}_1$.

   a) Show that $S(\beta) = \|\mathbf{X}\beta - \boldsymbol{y}\|_2^2$ can be decomposed, such that

$$S(\beta) = \|\mathbf{Q}_1^\top(\mathbf{X}\beta - \boldsymbol{y})\|_2^2 + \|\mathbf{Q}_2^\top \boldsymbol{y}\|_2^2$$

   Hint: As $\mathbf{Q}$ is in fact a rotation matrix, we have $\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_n$. $S(\beta)$ is unchanged under rotation, specifically $S(\beta) = \|\mathbf{X}\beta - \boldsymbol{y}\|_2^2 = \|\mathbf{Q}^\top(\mathbf{X}\beta - \boldsymbol{y})\|_2^2$. Use that $\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}_1\mathbf{Q}_1^\top + \mathbf{Q}_2\mathbf{Q}_2^\top$

   b) For $\hat\beta = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{y}$, show that $S(\hat\beta) = \|\mathbf{Q}_2^\top\boldsymbol{y}\|_2^2$

From b) we find that the usual estimator of σ^2 is given by

$$s^2 = \frac{S(\hat\beta)}{n-p} = \frac{\|\mathbf{Q}_2^\top\boldsymbol{y}\|_2^2}{n-p}.$$

Bates and Watts (1988) use the derived results in a) and b) to argue that (1) is an appropriate measure of convergence as it measures the orthogonality of the residuals to the expectation surface (its linear approximation) relative to the statistical uncertainty in the parameters (see Bates and Watts, 1988, pp. 50–51 in online version – link on moodle).

2. A rational function, $f\big(x,(\alpha,\beta)\big)$, is simply the ratio of two polynomial functions, $g(x,\alpha)$ and $h(x,\beta)$,

$$f\big(x,(\alpha,\beta)\big) = \frac{g(x,\alpha)}{h(x,\beta)} = \frac{\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_p x^p}{\beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_q x^q}, \tag{2}$$

where order of p and q need not to the same. To avoid over-parameterisation, we set $\beta_0 \equiv 1$ and rescale the other parameters accordingly.

Rational functions are quite versatile and can adapt to various behaviours of nonlinearity in the data.

However, for any given data set, f may have q roots in the support of x, implying vertical asymptotes – not likely to appear in the data.

Compared to most of the functions discussed in the course, it is rare that rational functions has a natural motivation as mean functions in nonlinear regression. Hence, we may think of them as *black box* functions.

a) Let $y_i = f\big(x_i,(\alpha,\beta)\big) + \varepsilon_i$, where f is given in (2). Outline a procedure for obtaining good starting values for α and β based on some data $(\boldsymbol{y}, \boldsymbol{x})$.

   Hint: Obtain a linear expression for y_i with x_i^k and $x_i^j y_i$, $k = 1,\ldots,p$ and $j = 1,\ldots q$ on the right hand side. For numerical stability it may be necesarry to standardise the variables, $\boldsymbol{y}$ and $\boldsymbol{x}$

b) Based on your procedure above, write a function for fitting a rational function to data using `nls`

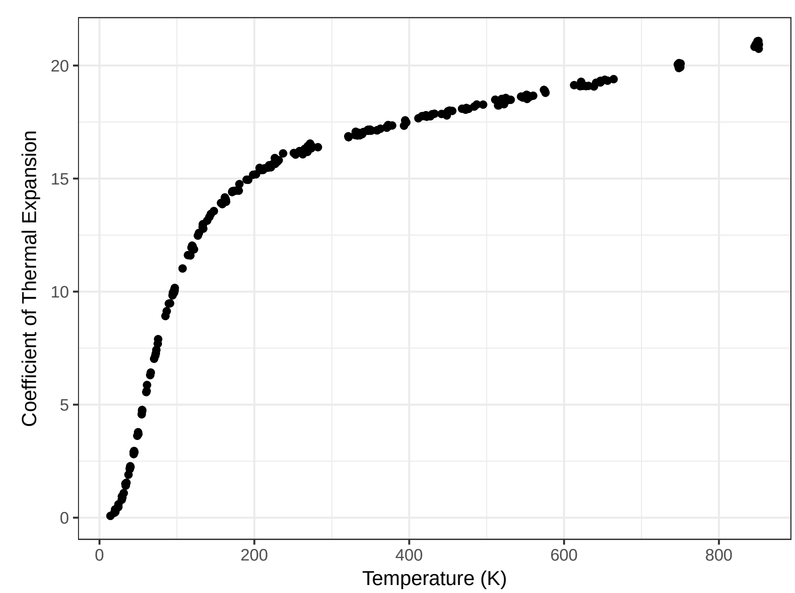

Figure 1: Thermal expansion of copper as function of temperature ($^\circ K$).

c) The `thermal` (on moodle, `thermal.csv`) data plotted in Figure 1 originates from an experiment where the aim was to study the thermal expansion of copper at different temperatures.

   Fit a rational function to the data using your procedure above. Start by fitting a quadratic-quadratic function, $p = q = 2$, and use the residuals to assess whether it is adequate. If not, increase the powers, refit and reassess the model fit.