“Linear” Regression With A Density Model

May 8, 2017

Let’s replace the standard straight-line model with a more general model.

The space we work in is 2-dimensional, $\boldsymbol{x} = [x_1, x_2]^T$. The standard straight-line model is $x_2 = mx_1 + c$. Viewed as a density, this is $\rho(\boldsymbol{x}) \propto \delta^2(x_2 - mx_1 - c)$.

We’ll replace this with a density that is uniform along the line $x_2 = mx_1 + c$, and falls away from that line with a Gaussian profile, so a contour map of $\rho(\boldsymbol{x})$ looks like a mountain range with a ridge along the “regression line”, and Gaussian slopes.

This is easy to accomplish. Let $\boldsymbol{m} \equiv [1, m]^T$ be the tangent vector to the line, so that $\boldsymbol{t} \equiv [-m, 1]^T$ is a vector satisfying $\boldsymbol{m}^T\boldsymbol{t} = 0$. Then,

$$\rho(\boldsymbol{x}) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{x}_0\right)^T \boldsymbol{A}\left(\boldsymbol{x} - \boldsymbol{x}_0\right)\right\},$$

with

$$\boldsymbol{A} \equiv s^{-2}\boldsymbol{t}\boldsymbol{t}^T$$

and

$$\boldsymbol{x}_0 = [0, c]^T.$$

The matrix $\boldsymbol{A}$ has a zero eigenvalue along $\boldsymbol{m}$ ($\boldsymbol{A}\boldsymbol{m} = \boldsymbol{0}$), so $\boldsymbol{A}$ is the “inverse covariance” (not really) of a covariance matrix with an infinite σ along the direction $\boldsymbol{m}$. Our line has thus been replaced with an infinitely-elongated Gaussian. The contours of constant $\rho(\boldsymbol{x})$ are all straight lines with tangent $\boldsymbol{m}$, that is, parallel to the regression line.

Suppose we have N observations $\boldsymbol{y}_i$, $i = 1, \ldots, N$, with measurement errors described by individual covariance matrices $\boldsymbol{C}_i$. The likelihood $\mathcal{L}_i$ for event i is then

$$\begin{aligned}
\mathcal{L}_i &= \int d^2\boldsymbol{x}\, \rho(\boldsymbol{x})\, \exp\left\{-\frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{y}_i\right)^T \boldsymbol{C}_i^{-1}\left(\boldsymbol{x} - \boldsymbol{y}_i\right)\right\} \\
&\propto \int d^2\boldsymbol{x}\, \exp\left\{-\frac{1}{2}\left[\left(\boldsymbol{x} - \boldsymbol{x}_0\right)^T \boldsymbol{A}\left(\boldsymbol{x} - \boldsymbol{x}_0\right) + \left(\boldsymbol{x} - \boldsymbol{y}_i\right)^T \boldsymbol{C}_i^{-1}\left(\boldsymbol{x} - \boldsymbol{y}_i\right)\right]\right\} \\
&= \int d^2\boldsymbol{u}\, \exp\left\{-\frac{1}{2}\left[\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} + \left(\boldsymbol{u} - \boldsymbol{y}_i + \boldsymbol{x}_{\boldsymbol{0}}\right)^T \boldsymbol{C}_i^{-1}\left(\boldsymbol{u} - \boldsymbol{y}_i + \boldsymbol{x}_0\right)\right]\right\}.
\end{aligned}$$

The argument of the exponential is $-(a/2)$ where

$$a \equiv \left(\boldsymbol{u} - \boldsymbol{u}_0\right)^T\left(\boldsymbol{C}_i^{-1} + \boldsymbol{A}\right)\left(\boldsymbol{u} - \boldsymbol{u}_0\right) + \left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)^T \boldsymbol{C}_i^{-1}\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right) - \boldsymbol{u}_0^T\left(\boldsymbol{C}_i^{-1} + \boldsymbol{A}\right)\boldsymbol{u}_0$$

where

$$\boldsymbol{u}_0 \equiv \left(C_i^{-1} + \boldsymbol{A}\right)^{-1} C_i^{-1} \left(\boldsymbol{y}_i - \boldsymbol{x}_0\right).$$

The integral may be performed straightforwardly. Up to immaterial constants, the result is

$$\mathcal{L}_i \propto \det\left[C_i^{-1} + \boldsymbol{A}\right]^{-1/2} \exp\left\{-\frac{1}{2}\left[\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)^T C_i^{-1} \left(\boldsymbol{y}_i - \boldsymbol{x}_0\right) - \boldsymbol{u}_0^T \left(C_i^{-1} + \boldsymbol{A}\right) \boldsymbol{u}_0\right]\right\}.$$

Substituting for $\boldsymbol{u}_0$ and $\boldsymbol{A}$, this is

$$\mathcal{L}_i \propto \det\left[C_i^{-1} + s^{-2}\boldsymbol{t}\boldsymbol{t}^T\right]^{-1/2} \exp\left\{-\frac{1}{2}\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)^T \left[C_i^{-1} - C_i^{-1}\left(C_i^{-1} + s^{-2}\boldsymbol{t}\boldsymbol{t}^T\right)^{-1} C_i^{-1}\right]\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)\right\}.$$

We may simplify this using the matrix inversion and determinant lemmas:

$$\begin{aligned}
\left(C_i^{-1} + s^{-2}\boldsymbol{t}\boldsymbol{t}^T\right)^{-1} &= C_i - C_i\boldsymbol{t}\left(s^2 + \boldsymbol{t}^T C_i \boldsymbol{t}\right)^{-1} \boldsymbol{t}^T C_i \\
&= C_i - \frac{C_i \boldsymbol{t}\boldsymbol{t}^T C_i}{s^2 + \boldsymbol{t}^T C_i \boldsymbol{t}},
\end{aligned}$$

and

$$\det\left[C_i^{-1} + s^{-2}\boldsymbol{t}\boldsymbol{t}^T\right] = \left[\det C_i\right]^{-1}\left(1 + s^{-2}\boldsymbol{t}^T C_i \boldsymbol{t}\right).$$

Exhibiting only terms containing some parameter dependence, we obtain for the log-likelihood

$$\begin{aligned}
\log \mathcal{L}_i &= -\frac{1}{2}\log\det\left(1 + s^{-2}\boldsymbol{t}^T C_i \boldsymbol{t}\right) - \frac{1}{2}\frac{\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)^T \boldsymbol{t}\boldsymbol{t}^T \left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)}{\left(s^2 + \boldsymbol{t}^T C_i \boldsymbol{t}\right)} \\
&= -\frac{1}{2}\log\det\left(1 + s^{-2}\boldsymbol{t}^T C_i \boldsymbol{t}\right) - \frac{1}{2}\frac{w^2}{\left(s^2 + \boldsymbol{t}^T C_i \boldsymbol{t}\right)},
\end{aligned}$$

where $w \equiv \boldsymbol{t}^T\left(\boldsymbol{y}_i - \boldsymbol{x}_0\right)$.

The full log-likelihood is then $\log \mathcal{L} = \sum_i \mathcal{L}_i$. This is a function of the model parameters m,c, s^2 that may be maximized, or traded in for a posterior density over the parameters, and fed to an MCMC.