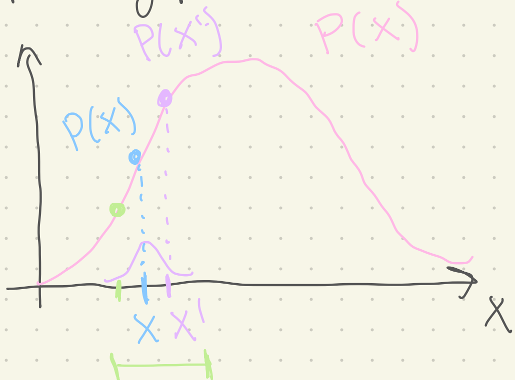
MH for linear fit

Nikki Rider
F2020

metropolis hastings (for linear fit)

aka finding a distribution for our parameters based on a probability function

last time:



this time:

$x \longrightarrow \theta = \begin{bmatrix} b \\ m \end{bmatrix}$

$\theta' = \theta + \delta * (2u-1)$

$u \in [0,1]$

$p \longrightarrow \mathcal{L} \equiv \prod_{i=1}^{N} p(y_i|x_i, \sigma_{yi}, \theta).$

$\delta = \begin{bmatrix} \delta_b \\ \delta_m \end{bmatrix}$

$$p(y_i|x_i, \sigma_{yi}, \theta) \equiv \frac{1}{\sqrt{2\pi\sigma_{yi}^2}} \exp\left(-\frac{(y_i - f(x_i, \theta))^2}{2\sigma_{yi}^2}\right).$$

data: $\{y_i\}, \{x_i\}, \{\sigma_{yi}\}$

model: $f(x_i, \theta) = b + m x_i$

MH: 2 tests

$$a \equiv \frac{P(x')}{P(x)}$$

$$u \in [0,1]$$

if $a > 1$ → keep x'

$x \to \theta$

$P(x) \to \mathcal{L}(\theta)$

else if $a > u$ → keep x'

else store old x again

likelihood → $\ln \mathcal{L}$

$$a = \frac{\mathcal{L}(\theta')}{\mathcal{L}(\theta)} > 1$$

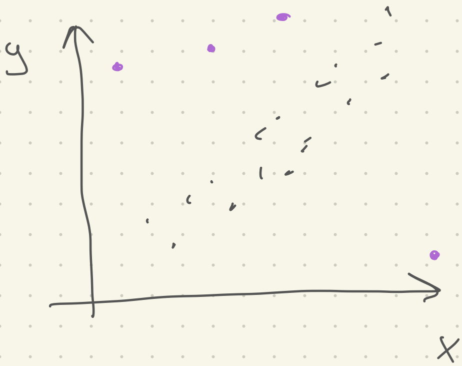
→ $\ln \mathcal{L}(\theta') - \ln \mathcal{L}(\theta) > 0$

$a > u$

→ $\ln \mathcal{L}(\theta') - \ln \mathcal{L}(\theta) > \ln(u)$

part D

$$\Theta = \begin{bmatrix} b \\ m \\ P_b \\ Y_b \\ V_b \end{bmatrix}$$



P_b : % of points that are outliers

Y_b : average " "

V_b : variance " "

$$\begin{aligned}
\mathcal{L} &\equiv p(\{y_i\}_{i=1}^{N} | m, b, P_b, Y_b, V_b) && (4) \\
&= \prod_{i=1}^{N} \left((1 - P_b) p_{good}(y_i | m, b) + P_b p_{bad}(y_i | Y_b, V_b) \right) && (5) \\
&\propto \prod_{i=1}^{N} \left(\frac{1 - P_b}{\sqrt{2\pi\sigma_{yi}^2}} \exp\left(-\frac{(y_i - m x_i - b)^2}{2\sigma_{yi}^2} \right) + \frac{P_b}{\sqrt{2\pi(V_b + \sigma_{yi}^2)}} \exp\left(-\frac{(y_i - Y_b)^2}{2(V_b + \sigma_{yi}^2)} \right) \right) . && (6)
\end{aligned}$$

"priors" = known restrictions

- $0 \le P_b \le 1$
- $V_b > 0$

prob. fcn = $\mathcal{L}(\Theta) * \text{prior}(\Theta)$

ln space → $\ln \mathcal{L}(\Theta) + \boxed{\ln \text{prior}(\Theta)}$

fcn prior

$$\Theta = \begin{bmatrix} b \\ m \end{bmatrix}$$

$$\mathscr{L} = \prod_{i=1}^{N} p(y_i | x_i, \sigma_{yi}, m, b)$$

$$p(y_i|x_i, \sigma_{yi}, m, b) = \frac{1}{\sqrt{2\,\pi\,\sigma_{yi}^2}} \exp\left(-\frac{[y_i - m\,x_i - b]^2}{2\,\sigma_{yi}^2}\right)$$

$$\ln \mathscr{L} = \ln \prod_{i=1}^{N} \underline{\frac{1}{\sqrt{2\pi\sigma_{y_i}^2}}}\ \underbrace{\exp\left(-\frac{(y_i - m x_i - b)^2}{2\sigma_{y_i}^2}\right)}$$

$$\ln \mathscr{L} = \sum_i \underbrace{\left[-\frac{1}{2}\ln\left(2\pi\sigma_{y_i}^2\right) - \frac{(y_i - m x_i - b)^2}{2\sigma_{y_i}^2}\right]}_{= \text{np.sum}\left(-\frac{1}{2}\,\text{np.log}(2*\text{np.pi}*\text{sigy}**2) - \frac{1}{2}(y - mx - b)\right.}$$

in code:

ln_likelihood( ) returns

$$\text{np.sum}\left(-\frac{1}{2}\,\text{np.log}(2\pi\ \text{sigy}**2) - \frac{1}{2}\left((y - \underset{\Theta[1]}{m}x - \underset{\Theta[0]}{b})/\text{sigy}\right)**2\right)$$