Metropolis with error Variance

- So far we have assumed

$$\vec{y} = f(\vec{q}; \vec{\theta}) + \vec{\varepsilon}, \quad \vec{\varepsilon} \sim N(0, \sigma_\varepsilon^2)$$

where σ_ε^2 is i) known or ii) estimated.

- we typically set $\sigma_\varepsilon^2 = \frac{1}{N_y - P} \sum_{i=1}^{N_y} \left(y_i - f(q_i; \vec{\theta}^0)\right)^2$

- However, we can also infer σ_ε^2 using a conjugate prior + MCMC

- If we assume $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ for our likelihood, then we can assume σ_ε^2 has an inverse-gamma prior

$$\pi_0(\sigma_\varepsilon^2) \propto (\sigma_\varepsilon^2)^{-(\alpha+1)} \cdot \exp\left(\beta / \sigma_\varepsilon^2\right) \sim \text{inv gamma}(\alpha, \beta)$$

- Because $\pi_0(\sigma_\varepsilon^2)$ is a conjugate prior,

$$\pi(\sigma_\varepsilon^2 \mid \vec{\theta}, \vec{y}) \propto (\sigma_\varepsilon^2)^{-(\alpha+1+N_y/2)} \exp\left(\frac{\beta + SS(\theta)/2}{\sigma_\varepsilon^2}\right)$$

$$\Rightarrow \pi(\sigma_\varepsilon^2 \mid \vec{\theta}, \vec{y}) \sim \text{inv. gamma}\left(\alpha + \frac{N_y}{2}, \beta + \frac{SS(\theta)}{2}\right)$$

Metropolis-Hastings (MH)

- The MH algorithm extends the metropolis algorithm beyond symmetric proposals, $J(\vec{\theta}^* \mid \vec{\theta}^{q-1})$.

- Idea: we need to account for $J(\vec{\theta}^* \mid \vec{\theta}^{q-1})$ explicitly in acceptance

$\pi(\vec{\theta}^{q-1} \mid \vec{y})$

- Idea: We need to account for

$$r_{MH} = \left(\frac{\pi(\vec{\theta}^* | \vec{y})}{J(\vec{\theta}^* | \vec{\theta}^{\ell-1})} \right) \Bigg/ \left(\frac{\pi(\vec{\theta}^{\ell-1} | \vec{y})}{J(\vec{\theta}^{\ell-1} | \vec{\theta}^*)} \right)$$

$$= \frac{p(\vec{y} | \vec{\theta}^*) \, \pi_0(\vec{\theta}^*) \, J(\vec{\theta}^{\ell-1} | \vec{\theta}^*)}{p(\vec{y} | \vec{\theta}^{\ell-1}) \, \pi_0(\vec{\theta}^{\ell-1}) \, J(\vec{\theta}^* | \vec{\theta}^{\ell-1})}$$

Note: If J is symmetric (think Gaussian)

$$J(\vec{\theta}^* | \vec{\theta}^{\ell-1}) \propto \exp\left(-\frac{1}{2}(\vec{\theta}^* - \vec{\theta}^{\ell-1})^T \underset{\sim}{V}^{-1} (\vec{\theta}^* - \vec{\theta}^{\ell-1})\right)$$

$$= J(\vec{\theta}^{\ell-1} | \vec{\theta}^*) \propto \exp\left(-\frac{1}{2}(\vec{\theta}^{\ell-1} - \vec{\theta}^*)^T \underset{\sim}{V}^{-1} (\vec{\theta}^{\ell-1} - \vec{\theta}^*)\right)$$

Stationarity

- Why does the chain converge to $\pi(\vec{\theta} | \vec{y})$?

i) Does the algorithm satisfy detailed balance?

- Note:
  - a chain with transition $\underset{\sim}{P}$ + distribution $\vec{\pi}$ is reversible if

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall x_i, x_j \in S$$

+ satisfies detailed balance.

- Detailed Balance $\Rightarrow$ Stationary distribution, $\vec{\pi}$.

- Let our stationary distribution be $\pi(\vec{\theta} | \vec{y})$.

  . Recall that the acceptance criteria for a symmetric proposal, $J(\vec{\theta}^* | \vec{\theta}^{\ell-1})$

- Let

- Recall that the acceptance criteria for a symmetric proposal, $J(\vec{\theta}^{*}|\vec{\theta}^{q-1})$ is given by

$$\rho(\theta^{*}, \vec{\theta}^{q-1}) = \min\left(\frac{\pi(\vec{\theta}^{*}|\vec{y})}{\pi(\vec{\theta}^{q-1}|\vec{y})}, 1\right) = \min\left(\frac{p(\vec{y}|\vec{\theta}^{*})\,\pi_0(\vec{\theta}^{*})}{p(\vec{y}|\vec{\theta}^{q-1})\,\pi_0(\vec{\theta}^{q-1})}, 1\right)$$

- Using this, the transition probability is

$$P_{*,q-1} = \begin{cases} \rho(\vec{\theta}^{*}, \vec{\theta}^{q-1})\, J(\vec{\theta}^{*}|\vec{\theta}^{q-1}), & \vec{\theta}^{*} \neq \vec{\theta}^{q-1} \\ & \end{cases}$$