STA 360/602L: Module 7.4

Metropolis within Gibbs

Dr. Olanrewaju Michael Akande

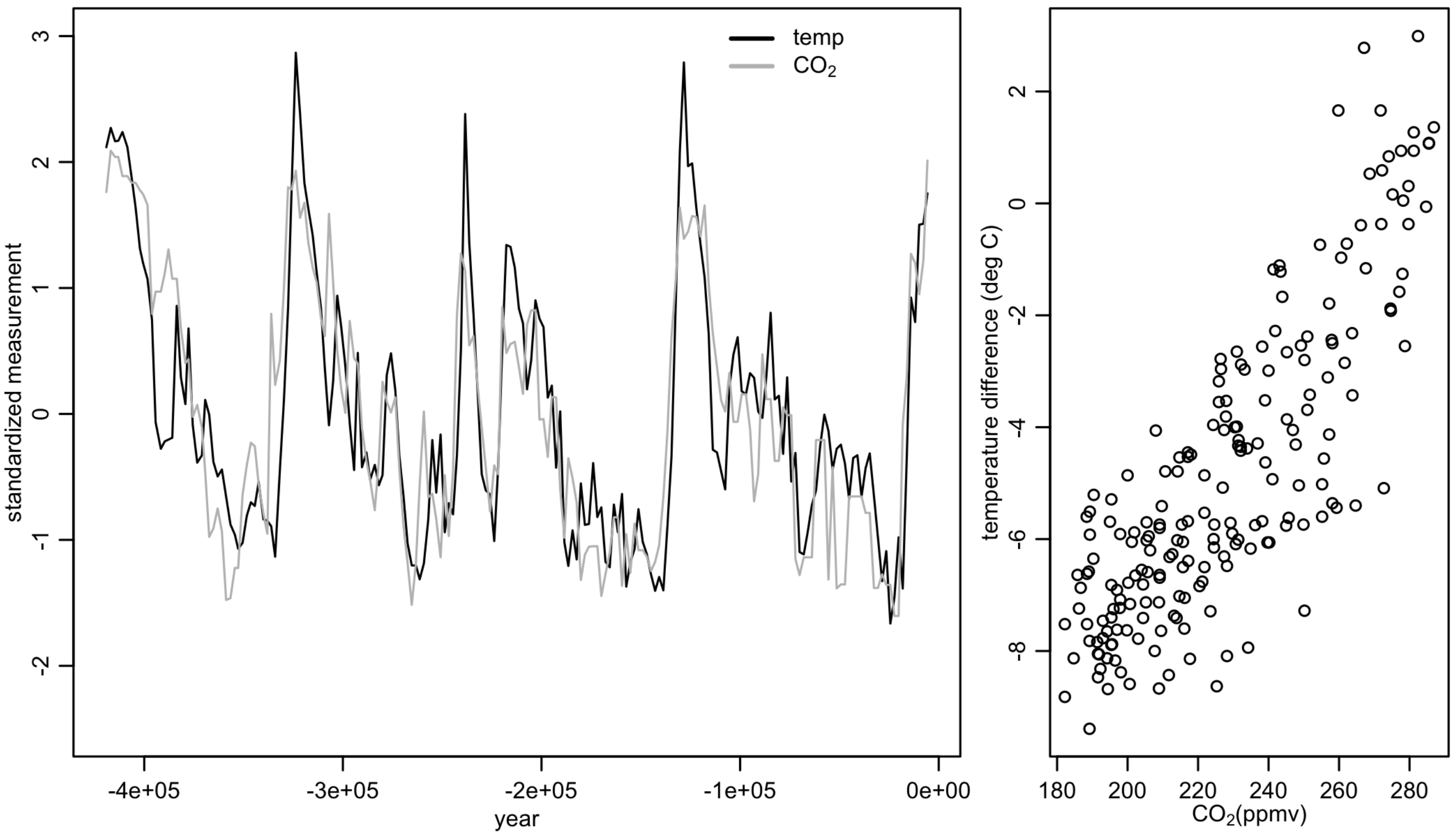
Combining Metropolis and Gibbs

- It is often the case that full conditionals are available for some parameters but not all.
- Very useful trick is to combine Gibbs and Metropolis.
- We will illustrate this by analyzing time series data on global warming.

Carbon dioxide and temperature

- Data are from analysis of ice cores from East Antarctica
- Temperature (recorded in terms of difference from current present temp in degrees C) and CO_2 (measured in ppm by volume) are standardized to have mean 0 and variance 1.
- 200 values, each roughly 2000 years apart.
- CO_2 values matched with temperature values roughly 1000 years later.

Data



CO_2 and temperature follow similar patterns over time.

Inference

- Interest lies in predicting temperature as a function of CO_2.
- In these data, the error terms are temporally correlated so that a reasonable model for temperature is

$$\boldsymbol{Y} \sim \mathcal{N}_n(\boldsymbol{X}\boldsymbol{\beta}, \Sigma),$$

  where $\boldsymbol{X}$ contains a column for the intercept plus a column for CO_2, and Σ has a first-order autoregressive structure so that:

$$\Sigma = \sigma^2 \boldsymbol{C}_\rho = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix}$$

- The covariance model assumes constant variance but a decreasing correlation as the time between temperature measures is greater.

Posterior inference

- We need to specify prior distributions for $\boldsymbol{\beta}$, σ^2 and ρ.
- If we assume

$$\pi(\boldsymbol{\beta}) = \mathcal{N}_p(\boldsymbol{\mu}_0, \Lambda_0),$$

then

$$\pi(\boldsymbol{\beta}|\boldsymbol{y}, \boldsymbol{X}, \sigma^2, \rho) = \mathcal{N}_p(\boldsymbol{\mu}_n, \Lambda_n),$$

where

$$\Lambda_n = \left[\Lambda_0^{-1} + \frac{1}{\sigma^2}\boldsymbol{X}^T \boldsymbol{C}_\rho^{-1} \boldsymbol{X}\right]^{-1}$$

$$\boldsymbol{\mu}_n = \Lambda_n \left[\Lambda_0^{-1}\boldsymbol{\mu}_0 + \frac{1}{\sigma^2}\boldsymbol{X}^T \boldsymbol{C}_\rho^{-1} \boldsymbol{y}\right].$$

Posterior inference

- If we assume

$$\pi(\sigma^2) = \mathcal{IG}\left(\frac{\nu_0}{2}, \frac{\nu_0\sigma_0^2}{2}\right),$$

then

$$\pi(\sigma^2|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\beta}, \rho) = \mathcal{IG}\left(\frac{\nu_n}{2}, \frac{\nu_n\sigma_n^2}{2}\right),$$

where

$$\nu_n = \nu_0 + n$$

$$\sigma_n^2 = \frac{1}{\nu_n}\left[\nu_0\sigma_0^2 + (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^T \boldsymbol{C}_\rho^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right] = \frac{1}{\nu_n}\left[\nu_0\sigma_0^2 + \mathrm{SSR}(\boldsymbol{\beta}, \rho)\right].$$

- Therefore, given ρ, we can use Gibbs sampling to cycle through the full conditionals for β and σ^2.

Posterior inference

- Next, we need a prior for the correlation ρ. There is no semi-conjugate option here.
- Since we expect ρ to be positive, we could use $\pi(\rho) = \text{Unif}(0, 1)$.
- Unfortunately, this does not lead to a standard full conditional.
- However, we can use Metropolis-Hastings for the resulting full conditional for ρ. Actually, if we could come up with a symmetric proposal for ρ, we can just use the Metropolis algorithm.
- So, technically, we have a Gibbs sampler since we will cycle through full conditionals. However, the sampling step for ρ will rely on Metropolis.
- Therefore, we have a Metropolis within Gibbs sampler.

Posterior inference

- Update for ρ (Metropolis) at iteration $(s+1)$:

  1. Generate a *candidate* value $\rho^\star \sim \text{Unif}(\rho^{(s)} - \delta, \rho^{(s)} + \delta)$. If $\rho^\star < 0$, reassign as $|\rho^\star|$. If $\rho^\star > 1$, reassign as $2 - \rho^\star$.

     I leave the proof that this "reflecting random walk" is symmetric to you.

  2. Compute the acceptance ratio

$$r = \frac{p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}^{(s+1)}, \sigma^{2(s+1)}, \rho^\star) \cdot \pi(\rho^\star)}{p(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\beta}^{(s+1)}, \sigma^{2(s+1)}, \rho^{(s)}) \cdot \pi(\rho^{(s)})}.$$

  3. Sample $u \sim U(0,1)$ independently and set

$$\rho^{(s+1)} = \begin{cases} \rho^\star & \text{if} \quad u < r \\ \rho^{(s)} & \text{if} \quad \text{otherwise} \end{cases}.$$

- So, for each iteration, we first sample from the full conditionals for β and σ^2, and then use this step to update ρ.

Move to the R script here.

What's next?

Move on to the readings for the next module!