Verification for the Gibbs Sampler

The fact that the Gibbs Sampler has the posterior density as an invariant distribution is not much more than the definition of conditional probability. We need some notation.

- $\Theta = (\theta_1, \theta_2, \ldots, \theta_d)$ is the parameter vector for the model.
- $\Pi(\Theta)$ a probability distribution on d-dimensional Euclidean space (in Bayesian applications Π is the posterior density of the parameters).
- $\Theta^{[-j]} = (\theta_1, \theta_2, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_d)$, the parameter vector with θ_j omitted.
- A full step of the Gibbs Sampler consists of d substeps in which a single element of Θ is updated. Let k be the substep index, with Θ_k denoting the value of Θ after the k^{th} substep, and p_k denoting the probability distribution of Θ_k.

A Gibbs substep updating θ_j consists of choosing $\theta_{j,k+1}$ according to its conditional distribution under Π given the rest of Θ at step k, that is

$$\theta_{j,k+1} \sim \Pi(\theta_j | \Theta^{[-j]} = \Theta_k^{[-j]}). \tag{1}$$

It suffices to show that Π is invariant under a substep. We have

$$\begin{aligned} p_{k+1}(\Theta) &= p_{k+1}(\theta_j | \Theta^{[-j]}) p_{k+1}(\Theta^{[-j]}) \\ &= p_{k+1}(\theta_j | \Theta^{[-j]}) p_k(\Theta^{[-j]}) \\ &= \Pi(\theta_j | \Theta^{[-j]}) p_k(\Theta^{[-j]}). \end{aligned} \tag{2}$$

The first line is the definition of conditional probability; the second is the fact that the substep leaves $\Theta^{[-j]}$ unchanged; the third is the definition of the Gibbs substep. Then if $p_k = \Pi$ we have

$$p_{k+1}(\Theta) = \Pi(\theta_j | \Theta^{[-j]}) \Pi(\Theta^{[-j]}) = \Pi(\Theta)$$

so Π is indeed invariant under a substep.

The hard part is showing that the posterior density is the unique invariant distribution, and that the chain converges to it. In Bayesian applications this is generally skipped. Instead one looks at simulation output and "confirms" that the chain converges to the same distribution regardless of where it is started. This "shows" that there is a unique invariant distribution, which therefore must be Π.