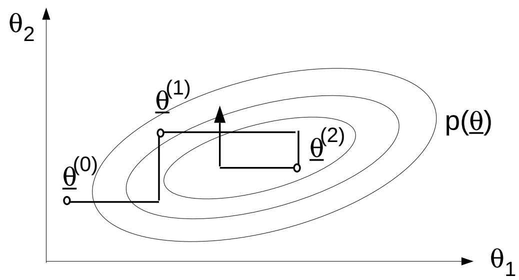
Gibbs sampling, continued



Sample $\theta_1^{(1)}$ from $p(\theta_1|\theta_2^{(0)}, \boldsymbol{y})$, then sample $\theta_2^{(1)}$ from $p(\theta_2|\theta_1^{(1)}, \boldsymbol{y})$, then repeat.

forms a Markov chain with stationary distribution .

For large n, $\boldsymbol{\theta}^{(n)}$ has distribution close to $p(\boldsymbol{\theta}|\boldsymbol{y})$, and more to the point, the sample $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(n)}$ 'looks like' a sample from $p(\boldsymbol{\theta}|\boldsymbol{y})$.

Example: Normal Random Sample

Suppose data $\boldsymbol{y} = (y_1, .., y_n)$ are a random sample from

Assume independent priors on μ and ϕ: ($\boldsymbol{\theta} = (\mu, \phi)$)

$$\sim \text{Normal}(\gamma, \kappa^{-1})$$

$$\sim \text{Gamma}(\alpha, \beta)$$

The posterior distribution $p(\mu, \phi|\boldsymbol{y}) =$ is (up to a constant of proportionality)

$$\phi^{n/2} \exp\left(-(\phi/2)\sum_{i=1}^{n}(y_i - \mu)^2\right) \qquad \phi^{\alpha-1}e^{-\beta\phi}$$

Gathering all terms which contain   gives

$$\propto \exp\left\{-\frac{\phi}{2}\left(n\ \ ^2 - \quad \sum y_i\right) - \frac{\kappa}{2}\left(\ \ ^2 - \gamma\ \ \right)\right\}$$

so that (completing the square)

$$\mu|\phi, \boldsymbol{y} \sim \text{Normal}\left(\frac{\phi\sum y_i + \gamma\kappa}{n\phi+\kappa}, \frac{1}{n\phi+\kappa}\right)$$

Doing the same for ϕ gives the full conditional distribution

$$\sim \text{Gamma}\left(\alpha + n/2, \beta + \sum (y_i - \mu)^2/2\right)$$

we implement the   by alternately drawing μ and ϕ from these distributions.

Systematic or random scan?

You do not have to cycle round the components   – e.g. you can choose which component to update next   .

Other variations?

You can update a   of several components at once, for example, draw (θ_2, θ_3) from $\pi(\quad|\theta_1, \theta_4, \ldots, \theta_k)$.

How does the Gibbs sampler work?

- Each step has   : we will spell this out later when we consider a generalisation of the Gibbs sampler.
- It is usually (i.e. except in pathological cases)   .

The BUGS program

Bayesian inference **U**sing **G**ibbs **S**ampling

- Language for specifying complex Bayesian models
- Constructs object-oriented internal representation of the model by identifying parents and children
- Builds up an complex model through specification of local structure
- Simulation from full conditionals using Gibbs sampling
- Current version (WinBUGS 1.4.1) runs in Windows, and incorporates the graphical model editor and a language for running in batch mode
- 'Classic' BUGS available for UNIX but this is an old version

WinBUGS is freely available from

`http://www.mrc-bsu.cam.ac.uk/bugs`

- An open source version of BUGS (called) is under development, and includes versions of BUGS that run under LINUX (LinBUGS) and that can be run directly from R (BRugs). See `http://www.rni.helsinki.fi/openbugs`

The Metropolis method

Sometimes it is to use the Gibbs sampler, e.g. when it is difficult to simulate a random value from a full conditional (e.g. if it is not a distribution of standard form – there's an example on slide).

In the Metropolis method, when the current state is $\boldsymbol{\theta}$, you draw a *proposed* new value from an distribution $q(\boldsymbol{\theta}'; \boldsymbol{\theta})$, that satisfies the symmetry condition $q(\boldsymbol{\theta}'; \boldsymbol{\theta}) =$.

Then you the new value with probability

$$= \min \left[1, \frac{\pi(\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})} \right]$$

and otherwise stay in the state $\boldsymbol{\theta}$.

Note that you *always* accept a proposed move to a *more probable* state.

Very commonly, the proposal is chosen so that, as in the Gibbs sampler, only component of $\boldsymbol{\theta}'$ is different from that of $\boldsymbol{\theta}$.

The Metropolis-Hastings method

You can generalise by the requirement that $q(\boldsymbol{\theta}'; \boldsymbol{\theta}) = q(\boldsymbol{\theta}; \boldsymbol{\theta}')$. Then the correct probability becomes

$$\alpha(\boldsymbol{\theta}'; \boldsymbol{\theta}) = \min \left[1, \frac{\pi(\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta})} \right]$$