Bayesian Methods in Genome Association Studies

Rohan L. Fernando

Iowa State University

February, 2010

Outline of Part I

Outline of Part II

Bayesian Regression Models

- Normal
- Student-t
- Mixture Models

Simulations

Part I

Bayesian Inference: Theory

Bayes Theorem

The conditional probability of X given Y is

$$\Pr(X|Y) = \frac{\Pr(X,Y)}{\Pr(Y)} = \frac{\Pr(Y|X)\Pr(X)}{\Pr(Y)}$$

where $\Pr(X,Y)$ is the joint probability of X and Y, $\Pr(X)$ is the probability of X, and $\Pr(Y)$ is the probability of Y.

Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

		Smoking: Yes	Smoking: No	
Lung Cancer	Yes	42,500	7,500	50,000
	No	207,500	742,500	950,000
		250,000	750,000	

Question: What is the relative frequency of lung cancer among smokers?

Answer: $\frac{42{,}500}{250{,}000} = 0.17$

Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

		Smoking: Yes	Smoking: No	
Lung Cancer	Yes	42,500	7,500	50,000
	No	207,500	742,500	950,000
		250,000	750,000	

Question: What is the relative frequency of lung cancer among smokers?

Answer: $\frac{42{,}500}{250{,}000} = 0.17$

Conditional Probability by Example

Joint distribution of smoking and lung cancer in a hypothetical population of 1,000,000:

		Smoking: Yes	Smoking: No	
Lung Cancer	Yes	42,500	7,500	50,000
	No	207,500	742,500	950,000
		250,000	750,000	

Question: What is the relative frequency of lung cancer among smokers?

Answer: $\frac{42{,}500}{250{,}000} = 0.17$

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Conditional Probability by Example

- As explained below, this relative frequency is also the conditional probability of lung cancer given smoking.
  - The frequentist definition of probability of an event is the limiting value of its relative frequency in a large number of trials.
  - Suppose we sample with replacement individuals from the 250,000 smokers and compute the relative frequency of lung cancer incidence.
  - It can be shown that as the sample size goes to infinity, this relative frequency will approach $\frac{42{,}500}{250{,}000} = 0.17$.
- This conditional probability is usually written as $\frac{42{,}500/1{,}000{,}000}{250{,}000/1{,}000{,}000} = 0.17$.
- The ratio in the numerator is joint probability of smoking and lung cancer, and the ratio in the denominator is the marginal probability of smoking.

Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency
- In Bayesian inference, probabilities are used to quantify your beliefs or knowledge about possible values of parameters
  - What is the probability that $h^2 > 0.5$?
  - What is the probability that milk yield is controlled by more than 100 loci?

Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency
- In Bayesian inference, probabilities are used to quantify your beliefs or knowledge about possible values of parameters
  - What is the probability that $h^2 > 0.5$?
  - What is the probability that milk yield is controlled by more than 100 loci?

Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency
- In Bayesian inference, probabilities are used to quantify your beliefs or knowledge about possible values of parameters
  - What is the probability that $h^2 > 0.5$?
  - What is the probability that milk yield is controlled by more than 100 loci?

Meaning of Probability in Bayesian Inference

- In the frequency approach, probability is a limiting frequency
- In Bayesian inference, probabilities are used to quantify your beliefs or knowledge about possible values of parameters
  - What is the probability that $h^2 > 0.5$?
  - What is the probability that milk yield is controlled by more than 100 loci?

Essentials of Bayesian Inference

- Prior probabilities quantify beliefs about parameters before the data are analyzed
- Parameters are related to the data through the model or "likelihood", which is the conditional probability density for the data given the parameters
- The prior and the likelihood are combined using Bayes theorem to obtain posterior probabilities, which are conditional probabilities for the parameters given the data
- Inferences about parameters are based on the posteior

Essentials of Bayesian Inference

- Prior probabilities quantify beliefs about parameters before the data are analyzed
- Parameters are related to the data through the model or "likelihood", which is the conditional probability density for the data given the parameters
- The prior and the likelihood are combined using Bayes theorem to obtain posterior probabilities, which are conditional probabilities for the parameters given the data
- Inferences about parameters are based on the posteior

Essentials of Bayesian Inference

- Prior probabilities quantify beliefs about parameters before the data are analyzed
- Parameters are related to the data through the model or "likelihood", which is the conditional probability density for the data given the parameters
- The prior and the likelihood are combined using Bayes theorem to obtain posterior probabilities, which are conditional probabilities for the parameters given the data
- Inferences about parameters are based on the posteior

Essentials of Bayesian Inference

- Prior probabilities quantify beliefs about parameters before the data are analyzed
- Parameters are related to the data through the model or "likelihood", which is the conditional probability density for the data given the parameters
- The prior and the likelihood are combined using Bayes theorem to obtain posterior probabilities, which are conditional probabilities for the parameters given the data
- Inferences about parameters are based on the posteior

Bayes Theorem in Bayesian Inference

- Let $f(\theta)$ denote the prior probability density for θ
- Let $f(\boldsymbol{y}|\theta)$ denote the likelihood
- Then, the posterior probability of θ is:

$$
\begin{aligned}
f(\theta|\boldsymbol{y}) &= \frac{f(\boldsymbol{y}|\theta)f(\theta)}{f(\boldsymbol{y})} \\
&\propto f(\boldsymbol{y}|\theta)f(\theta)
\end{aligned}
$$

Bayes Theorem in Bayesian Inference

- Let $f(\boldsymbol{\theta})$ denote the prior probability density for $\boldsymbol{\theta}$
- Let $f(\boldsymbol{y}|\boldsymbol{\theta})$ denote the likelihood
- Then, the posterior probability of θ is:

$$
\begin{aligned}
f(\boldsymbol{\theta}|\boldsymbol{y}) &= \frac{f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\boldsymbol{y})} \\
&\propto f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})
\end{aligned}
$$

Bayes Theorem in Bayesian Inference

- Let $f(\boldsymbol{\theta})$ denote the prior probability density for $\boldsymbol{\theta}$
- Let $f(\boldsymbol{y}|\boldsymbol{\theta})$ denote the likelihood
- Then, the posterior probability of $\boldsymbol{\theta}$ is:

$$
\begin{aligned}
f(\boldsymbol{\theta}|\boldsymbol{y}) &= \frac{f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}{f(\boldsymbol{y})} \\
&\propto f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})
\end{aligned}
$$

Computing posteriors

- Often no closed form for $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Further, even if computing $f(\boldsymbol{\theta}|\boldsymbol{y})$ is feasible, obtaining $f(\theta_i|\boldsymbol{y})$ would require integrating over many dimensions
- Thus, in many situations, inferences are made using the empirical posterior constructed by drawing samples from $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Gibbs sampler is widely used for drawing samples from posteriors

Computing posteriors

- Often no closed form for $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Further, even if computing $f(\boldsymbol{\theta}|\boldsymbol{y})$ is feasible, obtaining $f(\theta_i|\boldsymbol{y})$ would require integrating over many dimensions
- Thus, in many situations, inferences are made using the empirical posterior constructed by drawing samples from $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Gibbs sampler is widely used for drawing samples from posteriors

Computing posteriors

- Often no closed form for $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Further, even if computing $f(\boldsymbol{\theta}|\boldsymbol{y})$ is feasible, obtaining $f(\theta_i|\boldsymbol{y})$ would require integrating over many dimensions
- Thus, in many situations, inferences are made using the empirical posterior constructed by drawing samples from $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Gibbs sampler is widely used for drawing samples from posteriors

Computing posteriors

- Often no closed form for $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Further, even if computing $f(\boldsymbol{\theta}|\boldsymbol{y})$ is feasible, obtaining $f(\theta_i|\boldsymbol{y})$ would require integrating over many dimensions
- Thus, in many situations, inferences are made using the empirical posterior constructed by drawing samples from $f(\boldsymbol{\theta}|\boldsymbol{y})$
- Gibbs sampler is widely used for drawing samples from posteriors

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{lll} x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_2^t & \text{from} & f(x_2 | x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_3^t & \text{from} & f(x_3 | x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\ \vdots & & \vdots \\ x_n^t & \text{from} & f(x_n | x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t}) \end{array}$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{lll} x_1^t & \text{from} & f(x_1|x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_2^t & \text{from} & f(x_2|x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_3^t & \text{from} & f(x_3|x_1^t, x_2^t, \ldots, x_n^{t-1}) \\ \vdots & & \vdots \\ x_n^t & \text{from} & f(x_n|x_1^t, x_2^t, \ldots, x_{n-1}^t) \end{array}$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lll}
x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2 | x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3 | x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n | x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lll}
x_1^t & \text{from} & f(x_1|x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2|x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3|x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n|x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lll}
x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2 | x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3 | x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n | x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lll}
x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2 | x_1^t, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3 | x_1^t, x_2^t, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n | x_1^t, x_2^t, \ldots, x_{n-1}^t)
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lll}
x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2 | x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3 | x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n | x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$
\begin{array}{lcc}
x_1^t & \text{from} & f(x_1|x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2|x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3|x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n|x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}
$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{lcc}
x_1^t & \text{from} & f(x_1|x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_2^t & \text{from} & f(x_2|x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\
x_3^t & \text{from} & f(x_3|x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\
\vdots & & \vdots \\
x_n^t & \text{from} & f(x_n|x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t})
\end{array}$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Gibbs sampler

- Want to draw samples from $f(x_1, x_2, \ldots, x_n)$
- Even though it may be possible to compute $f(x_1, x_2, \ldots, x_n)$, it is difficult to draw samples directly from $f(x_1, x_2, \ldots, x_n)$
- Gibbs:
  - Get valid a starting point $\boldsymbol{x}^0$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{ccc} x_1^t & \text{from} & f(x_1 | x_2^{t-1}, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_2^t & \text{from} & f(x_2 | x_1^{t}, x_3^{t-1}, \ldots, x_n^{t-1}) \\ x_3^t & \text{from} & f(x_3 | x_1^{t}, x_2^{t}, \ldots, x_n^{t-1}) \\ \vdots & & \vdots \\ x_n^t & \text{from} & f(x_n | x_1^{t}, x_2^{t}, \ldots, x_{n-1}^{t}) \end{array}$$

- The sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ is a Markov chain with stationary distribution $f(x_1, x_2, \ldots, x_n)$

Inference from Markov chain

Can show that samples obtained from the Markov chain can be used to draw inferences from $f(x_1, x_2, \ldots, x_n)$ provided the chain is:

- Irreducible: can move from any state i to any other state j
- Positive recurrent: return time to any state has finite expectation
- *Markov Chains*, J. R. Norris (1997)

Inference from Markov chain

Can show that samples obtained from the Markov chain can be used to draw inferences from $f(x_1, x_2, \ldots, x_n)$ provided the chain is:

- Irreducible: can move from any state i to any other state j
- Positive recurrent: return time to any state has finite expectation
- *Markov Chains*, J. R. Norris (1997)

Inference from Markov chain

Can show that samples obtained from the Markov chain can be used to draw inferences from $f(x_1, x_2, \ldots, x_n)$ provided the chain is:

- Irreducible: can move from any state i to any other state j
- Positive recurrent: return time to any state has finite expectation
- *Markov Chains*, J. R. Norris (1997)

Inference from Markov chain

Can show that samples obtained from the Markov chain can be used to draw inferences from $f(x_1, x_2, \ldots, x_n)$ provided the chain is:

- Irreducible: can move from any state i to any other state j
- Positive recurrent: return time to any state has finite expectation
- *Markov Chains*, J. R. Norris (1997)

Example

Let $f(\boldsymbol{x})$ be a bivariate normal density with means

$$\mu' = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

and covariance matrix

$$\boldsymbol{V} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2.0 \end{bmatrix}$$

Suppose we do not know how to draw samples from $f(\boldsymbol{x})$, but know how to draw samples from $f(x_i|x_j)$, which is univariate normal with mean:

$$\mu_{i.j} = \mu_i + \frac{v_{ij}}{v_{jj}}(x_j - \mu_j)$$

and variance

$$v_{i.j} = v_{ii} - \frac{v_{ij}^2}{v_{jj}}$$

Gibbs sampler

- Gibbs:
  - Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{ccc} x_1^t & \text{from} & f(x_1 | x_2^{t-1}) \\ x_2^t & \text{from} & f(x_2 | x_1^t) \end{array}$$

- Use the sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ to compute any property of $f(\boldsymbol{x})$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$

Gibbs sampler

- Gibbs:
  - Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{ccc} x_1^t & \text{from} & f(x_1|x_2^{t-1}) \\ x_2^t & \text{from} & f(x_2|x_1^t) \end{array}$$

- Use the sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ to compute any property of $f(\boldsymbol{x})$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$

Gibbs sampler

- Gibbs:
  - Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{lll} x_1^t & \text{from} & f(x_1 | x_2^{t-1}) \\ x_2^t & \text{from} & f(x_2 | x_1^t) \end{array}$$

- Use the sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ to compute any property of $f(\boldsymbol{x})$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$

Gibbs sampler

- Gibbs:
  - Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{ccc} x_1^t & \text{from} & f(x_1|x_2^{t-1}) \\ x_2^t & \text{from} & f(x_2|x_1^t) \end{array}$$

- Use the sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ to compute any property of $f(\boldsymbol{x})$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$

Gibbs sampler

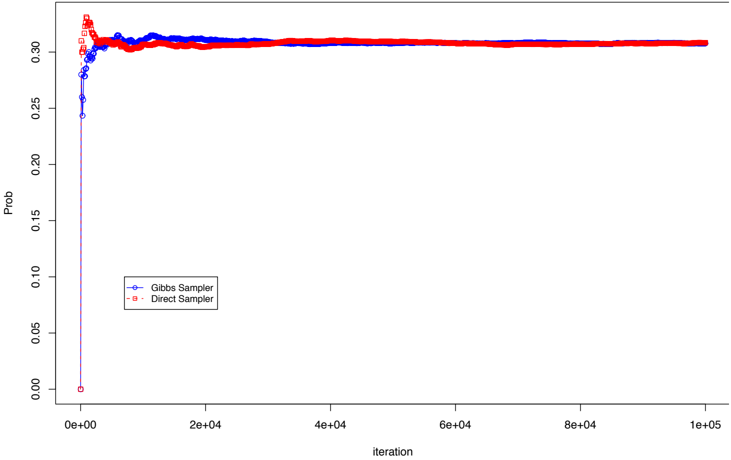
- Gibbs:
  - Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
  - Draw sample $\boldsymbol{x}^t$ as:

$$\begin{array}{ccc} x_1^t & \text{from} & f(x_1|x_2^{t-1}) \\ x_2^t & \text{from} & f(x_2|x_1^t) \end{array}$$

- Use the sequence $\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^n$ to compute any property of $f(\boldsymbol{x})$, for example

$$\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$$

MCMC Estimates of $\Pr(x_1 > \mu_1$ and $x_2 > \mu_2)$



Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min\left(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})}\right)$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min\left(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})}\right)$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Metropolis-Hastings sampler

- Sometimes may not be able to draw samples directly from $f(x_i|\boldsymbol{x}_{i\_})$
- Convergence of the Gibbs sampler may be too slow
- Metropolis-Hastings (MH) for sampling from $f(x)$:
  - a candidate sample, y, is drawn from a proposal distribution $q(y|x^{t-1})$
  - $$x^t = \begin{cases} y & \text{with probability } \alpha \\ x^{t-1} & \text{with probability } 1-\alpha \end{cases}$$
  - $$\alpha = \min(1, \frac{f(y)q(x^{t-1}|y)}{f(x^{t-1})q(y|x^{t-1})})$$
- The samples from MH is a Markov chain with stationary distribution $f(x)$

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

Proposal distributions

Two main types:

- Approximations of the target density: $f(x)$
  - Not easy to find approximation that is easy to sample from
  - High acceptance rate is good!
- Random walk type: stay close to the previous sample
  - Generally easy to construct proposal
  - High acceptance rate may indicate that candidate is too close to previous sample
  - Intermediate acceptance rate is good

MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $\boldsymbol{x}^t$ as:

$$\begin{aligned} y_1 &= x_1^{t-1} + u_1 \\ y_2 &= x_2^{t-1} + u_2 \end{aligned}$$

  where u_i is Uniform$(-v_{ii}^{1/2}, v_{ii}^{1/2})$.
- Compute

$$\alpha = min(1, \frac{f(\boldsymbol{y})}{f(\boldsymbol{x}^{t-1})})$$

  and

$$\boldsymbol{x}^t = \begin{cases} \boldsymbol{y} & \text{with probability } \alpha \\ \boldsymbol{x}^{t-1} & \text{with probability } 1-\alpha \end{cases}$$

MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $\boldsymbol{x}^t$ as:

$$\begin{aligned} y_1 &= x_1^{t-1} + u_1 \\ y_2 &= x_2^{t-1} + u_2 \end{aligned}$$

where u_i is Uniform$(-v_{ii}^{1/2}, v_{ii}^{1/2})$.

- Compute

$$\alpha = min(1, \frac{f(\boldsymbol{y})}{f(\boldsymbol{x}^{t-1})})$$

and

$$\boldsymbol{x}^t = \begin{cases} \boldsymbol{y} & \text{with probability } \alpha \\ \boldsymbol{x}^{t-1} & \text{with probability } 1-\alpha \end{cases}$$

MH Sampler to Estimate $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$

MH Sampler:

- Start with $\boldsymbol{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Draw sample $\boldsymbol{x}^t$ as:

$$\begin{aligned} y_1 &= x_1^{t-1} + u_1 \\ y_2 &= x_2^{t-1} + u_2 \end{aligned}$$

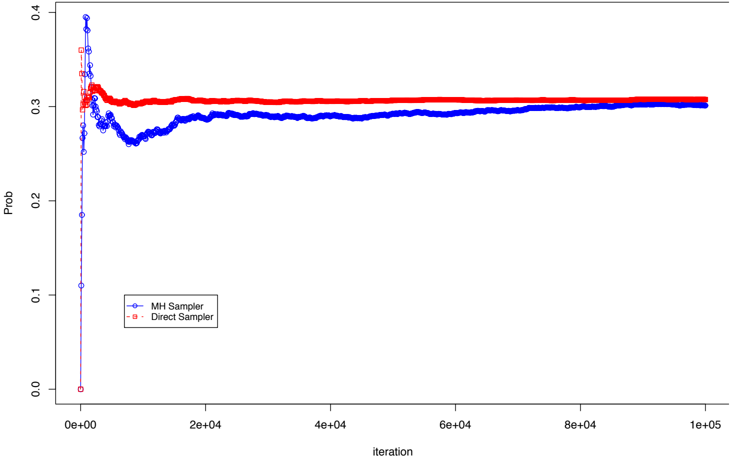
  where u_i is Uniform$(-v_{ii}^{1/2}, v_{ii}^{1/2})$.
- Compute

$$\alpha = min(1, \frac{f(\boldsymbol{y})}{f(\boldsymbol{x}^{t-1})})$$

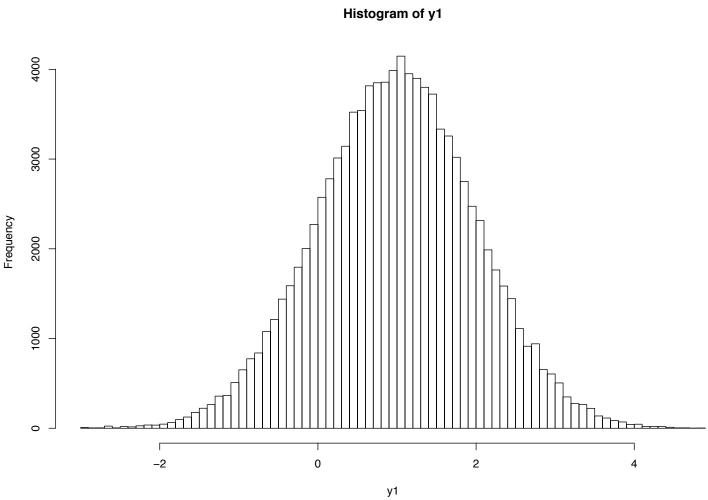
  and

$$\boldsymbol{x}^t = \begin{cases} \boldsymbol{y} & \text{with probability } \alpha \\ \boldsymbol{x}^{t-1} & \text{with probability } 1-\alpha \end{cases}$$

MCMC Estimates of $\Pr(x_1 > \mu_1 \text{ and } x_2 > \mu_2)$



Distribution of y_1 Sampled Using MH



Part II

Bayesian Inference: Application to Whole Genome Analyses

Model

Model:

$$y_i = \mu + \sum_j X_{ij}\alpha_j + e_i$$

Priors:

- $\mu \propto$ constant (not proper, but posterior is proper)
- $(e_i|\sigma_e^2) \sim (\text{iid})\text{N}(0, \sigma_e^2)$; $\sigma_e^2 \sim \nu_e S_e^2 \chi_{\nu_e}^{-2}$
- Consider several different priors for α_j

Model

Model:

$$y_i = \mu + \sum_j X_{ij}\alpha_j + e_i$$

Priors:

- $\mu \propto$ constant (not proper, but posterior is proper)
- $(e_i|\sigma_e^2) \sim (\text{iid})\text{N}(0, \sigma_e^2)$; $\sigma_e^2 \sim \nu_e S_e^2 \chi_{\nu_e}^{-2}$
- Consider several different priors for α_j

Model

Model:

$$y_i = \mu + \sum_j X_{ij}\alpha_j + e_i$$

Priors:

- $\mu \propto$ constant (not proper, but posterior is proper)
- $(e_i|\sigma_e^2) \sim$ (iid)N$(0, \sigma_e^2)$; $\sigma_e^2 \sim \nu_e S_e^2 \chi_{\nu_e}^{-2}$
- Consider several different priors for α_j

Model

Model:

$$y_i = \mu + \sum_j X_{ij}\alpha_j + e_i$$

Priors:

- $\mu \propto$ constant (not proper, but posterior is proper)
- $(e_i|\sigma_e^2) \sim$ (iid)N$(0, \sigma_e^2)$; $\sigma_e^2 \sim \nu_e S_e^2 \chi_{\nu_e}^{-2}$
- Consider several different priors for α_j

Normal

- Prior: $(\alpha_j|\sigma_\alpha^2) \sim$ (iid)N$(0, \sigma_\alpha^2)$; σ_α^2 is known
- What is σ_α^2?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}_i' \boldsymbol{\alpha}$$

  - Note that α is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}_i'$ being random
- So, σ_α^2 is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma_\alpha^2 = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}'_i\boldsymbol{\alpha}$$

  - Note that α is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma_\alpha^2) \sim$ (iid)N$(0, \sigma_\alpha^2)$; σ_α^2 is known
- What is σ_α^2?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}_i'\alpha$$

  - Note that α is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}_i'$ being random
- So, σ_α^2 is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma_\alpha^2 = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of *i* can be written as:

$$g_i = \mu + \boldsymbol{x}'_i \boldsymbol{\alpha}$$

  - Note that α is common to all *i*
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus *j* is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \mathrm{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}'_i\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}'_i\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}'_i\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma_\alpha^2) \sim \text{(iid)N}(0, \sigma_\alpha^2)$; σ_α^2 is known
- What is σ_α^2?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}_i'\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}_i'$ being random
- So, σ_α^2 is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma_\alpha^2 = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}_i'\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}_i'$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Normal

- Prior: $(\alpha_j|\sigma^2_\alpha) \sim$ (iid)N$(0, \sigma^2_\alpha)$; σ^2_α is known
- What is σ^2_α?
- Assume the QTL genotypes are a subset of those available for the analysis
  - Then, the genotypic value of i can be written as:

$$g_i = \mu + \boldsymbol{x}'_i\boldsymbol{\alpha}$$

  - Note that $\boldsymbol{\alpha}$ is common to all i
  - Thus, the variance of g_i comes from $\boldsymbol{x}'_i$ being random
- So, σ^2_α is not the genetic variance at a locus
- If locus j is randomly sampled from all the loci available for analysis:
  - Then, α_j will be a random variable
  - $\sigma^2_\alpha = \text{Var}(\alpha_j)$

Relationship of σ_α^2 to genetic variance

Assume loci with effect on trait are in linkage equilibrium. Then, the additive genetic variance is

$$V_A = \sum_j^k 2p_jq_j\alpha_j^2,$$

where $p_j = 1 - q_j$ is gene frequency at SNP locus j.
Letting $U_j = 2p_jq_j$ and $V_j = \alpha_j^2$,

$$V_A = \sum_j^k U_jV_j$$

For a randomly sampled locus, covariance between U_j and V_j is

$$C_{UV} = \frac{\sum_j U_jV_j}{k} - (\frac{\sum_j U_j}{k})(\frac{\sum_j V_j}{k})$$

Relationship of σ^2_α to genetic variance

Assume loci with effect on trait are in linkage equilibrium. Then, the additive genetic variance is

$$V_A = \sum_j^k 2p_j q_j \alpha_j^2,$$

where $p_j = 1 - q_j$ is gene frequency at SNP locus j.
Letting $U_j = 2p_j q_j$ and $V_j = \alpha_j^2$,

$$V_A = \sum_j^k U_j V_j$$

For a randomly sampled locus, covariance between U_j and V_j is

$$C_{UV} = \frac{\sum_j U_j V_j}{k} - \left(\frac{\sum_j U_j}{k}\right)\left(\frac{\sum_j V_j}{k}\right)$$

Relationship of σ_α^2 to genetic variance

Assume loci with effect on trait are in linkage equilibrium. Then, the additive genetic variance is

$$V_A = \sum_j^k 2p_j q_j \alpha_j^2,$$

where $p_j = 1 - q_j$ is gene frequency at SNP locus j.
Letting $U_j = 2p_j q_j$ and $V_j = \alpha_j^2$,

$$V_A = \sum_j^k U_j V_j$$

For a randomly sampled locus, covariance between U_j and V_j is

$$C_{UV} = \frac{\sum_j U_j V_j}{k} - \left(\frac{\sum_j U_j}{k}\right)\left(\frac{\sum_j V_j}{k}\right)$$

Relationship of σ_α^2 to genetic variance

Rearranging the previous expression for C_{UV} gives

$$\sum_j U_j V_j = kC_{UV} + (\sum_j U_j)(\frac{\sum_j V_j}{k})$$

So,

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)(\frac{\sum_j \alpha_j^2}{k})$$

Letting $\sigma_\alpha^2 = \frac{\sum_j \alpha_j^2}{k}$ gives

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)\sigma_\alpha^2$$

and,

$$\sigma_\alpha^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$

Relationship of σ_α^2 to genetic variance

Rearranging the previous expression for C_{UV} gives

$$\sum_j U_j V_j = kC_{UV} + (\sum_j U_j)(\frac{\sum_j V_j}{k})$$

So,

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)(\frac{\sum_j \alpha_j^2}{k})$$

Letting $\sigma_\alpha^2 = \frac{\sum_j \alpha_j^2}{k}$ gives

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)\sigma_\alpha^2$$

and,

$$\sigma_\alpha^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$

Relationship of σ_α^2 to genetic variance

Rearranging the previous expression for C_{UV} gives

$$\sum_j U_j V_j = kC_{UV} + (\sum_j U_j)(\frac{\sum_j V_j}{k})$$

So,

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)(\frac{\sum_j \alpha_j^2}{k})$$

Letting $\sigma_\alpha^2 = \frac{\sum_j \alpha_j^2}{k}$ gives

$$V_A = kC_{UV} + (\sum_j 2p_j q_j)\sigma_\alpha^2$$

and,

$$\sigma_\alpha^2 = \frac{V_A - kC_{UV}}{\sum_j 2p_j q_j}$$

Blocked Gibbs sampler

- Let $\boldsymbol{\theta}' = [\mu, \boldsymbol{\alpha}']$
- Can show that $(\theta|\boldsymbol{y}, \sigma_e^2) \sim \mathrm{N}(\hat{\theta}, \boldsymbol{C}^{-1}\sigma_e^2)$
- $$\hat{\theta} = \boldsymbol{C}^{-1}\boldsymbol{W}'\boldsymbol{y}; \quad \boldsymbol{W} = [\mathbf{1}, \boldsymbol{X}]$$
- $$\boldsymbol{C} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\boldsymbol{X} \\ \boldsymbol{X}'\mathbf{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$
- Blocked Gibbs sampler
  - García-Cortés and Sorensen (1996, GSE 28:121-126)
  - *Likelihood, Bayesian and MCMC Methods* $\cdots$ (LBMMQG, Sorensen and Gianola, 2002)

Blocked Gibbs sampler

- Let $\boldsymbol{\theta}' = [\mu, \boldsymbol{\alpha}']$
- Can show that $(\boldsymbol{\theta}|\boldsymbol{y}, \sigma_e^2) \sim \mathrm{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{C}^{-1}\sigma_e^2)$
- $$\hat{\theta} = \boldsymbol{C}^{-1}\boldsymbol{W}'\boldsymbol{y}; \quad \boldsymbol{W} = [\boldsymbol{1}, \boldsymbol{X}]$$
- $$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{1}'\boldsymbol{1} & \boldsymbol{1}'\boldsymbol{X} \\ \boldsymbol{X}'\boldsymbol{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$
- Blocked Gibbs sampler
  - García-Cortés and Sorensen (1996, GSE 28:121-126)
  - *Likelihood, Bayesian and MCMC Methods* $\cdots$ (LBMMQG, Sorensen and Gianola, 2002)

Blocked Gibbs sampler

- Let $\boldsymbol{\theta}' = [\mu, \boldsymbol{\alpha}']$
- Can show that $(\boldsymbol{\theta}|\boldsymbol{y}, \sigma_e^2) \sim \mathrm{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{C}^{-1}\sigma_e^2)$
- $$\hat{\boldsymbol{\theta}} = \boldsymbol{C}^{-1}\boldsymbol{W}'\boldsymbol{y}; \quad \boldsymbol{W} = [\mathbf{1}, \boldsymbol{X}]$$
- $$\boldsymbol{C} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\boldsymbol{X} \\ \boldsymbol{X}'\mathbf{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$
- Blocked Gibbs sampler
  - García-Cortés and Sorensen (1996, GSE 28:121-126)
  - *Likelihood, Bayesian and MCMC Methods* $\cdots$ (LBMMQG, Sorensen and Gianola, 2002)

Blocked Gibbs sampler

- Let $\boldsymbol{\theta}' = [\mu, \boldsymbol{\alpha}']$
- Can show that $(\boldsymbol{\theta}|\boldsymbol{y}, \sigma_e^2) \sim \mathrm{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{C}^{-1}\sigma_e^2)$
- $$\hat{\boldsymbol{\theta}} = \boldsymbol{C}^{-1}\boldsymbol{W}'\boldsymbol{y}; \quad \boldsymbol{W} = [\mathbf{1}, \boldsymbol{X}]$$
- $$\boldsymbol{C} = \begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\boldsymbol{X} \\ \boldsymbol{X}'\mathbf{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$
- Blocked Gibbs sampler
  - García-Cortés and Sorensen (1996, GSE 28:121-126)
  - *Likelihood, Bayesian and MCMC Methods* ⋯ (LBMMQG, Sorensen and Gianola, 2002)

Blocked Gibbs sampler

- Let $\boldsymbol{\theta}' = [\mu, \boldsymbol{\alpha}']$
- Can show that $(\boldsymbol{\theta}|\boldsymbol{y}, \sigma_e^2) \sim \mathrm{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{C}^{-1}\sigma_e^2)$
- $$\hat{\boldsymbol{\theta}} = \boldsymbol{C}^{-1}\boldsymbol{W}'\boldsymbol{y}; \quad \boldsymbol{W} = [\boldsymbol{1}, \boldsymbol{X}]$$
- $$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{1}'\boldsymbol{1} & \boldsymbol{1}'\boldsymbol{X} \\ \boldsymbol{X}'\boldsymbol{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix}$$
- Blocked Gibbs sampler
  - García-Cortés and Sorensen (1996, GSE 28:121-126)
  - *Likelihood, Bayesian and MCMC Methods* $\cdots$ (LBMMQG, Sorensen and Gianola, 2002)

Full conditionals for single-site Gibbs

- $(\mu|\boldsymbol{y},\boldsymbol{\alpha},\sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y}-\boldsymbol{X\alpha})}{n},\frac{\sigma_e^2}{n})$
- $(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j,\frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j'\neq j} \boldsymbol{x}_{j'}\alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \sim [(\boldsymbol{y}-\boldsymbol{W\theta})'(\boldsymbol{y}-\boldsymbol{W\theta})+\nu_e S_e^2]\chi^{-2}_{(\nu_e+n)}$

Full conditionals for single-site Gibbs

- $(\mu|\boldsymbol{y},\boldsymbol{\alpha},\sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y}-\boldsymbol{X\alpha})}{n}, \frac{\sigma_e^2}{n})$
- $(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j'\neq j} \boldsymbol{x}_{j'}\alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2|\boldsymbol{y},\mu,\alpha) \sim [(\boldsymbol{y}-\boldsymbol{W\theta})'(\boldsymbol{y}-\boldsymbol{W\theta}) + \nu_e S_e^2]\chi^{-2}_{(\nu_e+n)}$

Full conditionals for single-site Gibbs

- $(\mu|\boldsymbol{y},\boldsymbol{\alpha},\sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\alpha})}{n}, \frac{\sigma_e^2}{n})$
- $(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j'\neq j}\boldsymbol{x}_{j'}\alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \sim [(\boldsymbol{y}-\boldsymbol{W\theta})'(\boldsymbol{y}-\boldsymbol{W\theta}) + \nu_e S_e^2]\chi^{-2}_{(\nu_e+n)}$

Full conditionals for single-site Gibbs

- $(\mu|\boldsymbol{y}, \boldsymbol{\alpha}, \sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\alpha})}{n}, \frac{\sigma_e^2}{n})$
- $(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}) \sim [(\boldsymbol{y} - \boldsymbol{W}\boldsymbol{\theta})'(\boldsymbol{y} - \boldsymbol{W}\boldsymbol{\theta}) + \nu_e S_e^2]\chi^{-2}_{(\nu_e+n)}$

Full conditionals for single-site Gibbs

- $(\mu | \boldsymbol{y}, \boldsymbol{\alpha}, \sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\alpha})}{n}, \frac{\sigma_e^2}{n})$
- $(\alpha_j | \boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j' \boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'} \alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j' \boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2 | \boldsymbol{y}, \mu, \alpha) \sim [(\boldsymbol{y} - \boldsymbol{W}\boldsymbol{\theta})'(\boldsymbol{y} - \boldsymbol{W}\boldsymbol{\theta}) + \nu_e S_e^2] \chi^{-2}_{(\nu_e + n)}$

Full conditionals for single-site Gibbs

- $(\mu|\boldsymbol{y}, \boldsymbol{\alpha}, \sigma_e^2) \sim \mathrm{N}(\frac{\mathbf{1}'(\boldsymbol{y}-\boldsymbol{X\alpha})}{n}, \frac{\sigma_e^2}{n})$
- $(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \sim \mathrm{N}(\hat{\alpha}_j, \frac{\sigma_e^2}{c_j})$
  - $$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$
  - $$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'}$$
  - $$c_j = (\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2})$$
- $(\sigma_e^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}) \sim [(\boldsymbol{y} - \boldsymbol{W\theta})'(\boldsymbol{y} - \boldsymbol{W\theta}) + \nu_e S_e^2]\chi^{-2}_{(\nu_e+n)}$

Derive: full conditional for α_j

From Bayes' Theorem,

$$f(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) = \frac{f(\alpha_j,\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}{f(\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}$$

$$\propto f(\boldsymbol{y}|\alpha_j,\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)f(\alpha_j)f(\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2}\exp\{-\frac{(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}(\sigma_\alpha^2)^{-1/2}\exp\{-\frac{\alpha_j^2}{2\sigma_\alpha^2}\}$$

where

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j\neq j'}\boldsymbol{x}_{j'}\alpha_{j'}$$

Derive: full conditional for α_j

From Bayes' Theorem,

$$f(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) = \frac{f(\alpha_j, \boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2)}{f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2)}$$

$$\propto f(\boldsymbol{y}|\alpha_j, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) f(\alpha_j) f(\mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2} \exp\{-\frac{(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)'(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\} (\sigma_\alpha^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_\alpha^2}\}$$

where

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j \neq j'} \boldsymbol{x}_{j'}\alpha_{j'}$$

Derive: full conditional for α_j

From Bayes' Theorem,

$$f(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) = \frac{f(\alpha_j,\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}{f(\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}$$

$$\propto f(\boldsymbol{y}|\alpha_j,\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)f(\alpha_j)f(\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2}\exp\{-\frac{(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}(\sigma_\alpha^2)^{-1/2}\exp\{-\frac{\alpha_j^2}{2\sigma_\alpha^2}\}$$

where

$$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j\neq j'}\boldsymbol{x}_{j'}\alpha_{j'}$$

Derive: full conditional for α_j

From Bayes' Theorem,

$$f(\alpha_j|\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2) = \frac{f(\alpha_j,\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}{f(\boldsymbol{y},\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)}$$

$$\propto f(\boldsymbol{y}|\alpha_j,\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)f(\alpha_j)f(\mu,\boldsymbol{\alpha}_{j\_},\sigma_e^2)$$

$$\propto (\sigma_e^2)^{-n/2}\exp\{-\frac{(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}(\sigma_\alpha^2)^{-1/2}\exp\{-\frac{\alpha_j^2}{2\sigma_\alpha^2}\}$$

where

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j\neq j'}\boldsymbol{x}_{j'}\alpha_{j'}$$

Derive: full conditional for α_j

The exponential terms in the joint density can be written as:

$$-\frac{1}{2\sigma_e^2}\{\boldsymbol{w}'\boldsymbol{w} - 2\boldsymbol{x}_j'\boldsymbol{w}\alpha_j + [\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2\}$$

Completing the square in this expression with respect to α_j gives

$$-\frac{1}{2\sigma_e^2}\{c_j(\alpha_j - \hat{\alpha}_j)^2 + \boldsymbol{w}'\boldsymbol{w} - c_j\hat{\alpha}_j{}^2\}$$

where

$$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$

So,

$$f(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \propto \exp\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}}\}$$

Derive: full conditional for α_j

The exponential terms in the joint density can be written as:

$$-\frac{1}{2\sigma_e^2}\{\boldsymbol{w}'\boldsymbol{w} - 2\boldsymbol{x}_j'\boldsymbol{w}\alpha_j + [\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2\}$$

Completing the square in this expression with respect to α_j gives

$$-\frac{1}{2\sigma_e^2}\{c_j(\alpha_j - \hat{\alpha}_j)^2 + \boldsymbol{w}'\boldsymbol{w} - c_j\hat{\alpha}_j{}^2\}$$

where

$$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$

So,

$$f(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \propto \exp\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}}\}$$

Derive: full conditional for α_j

The exponential terms in the joint density can be written as:

$$-\frac{1}{2\sigma_e^2}\{\boldsymbol{w}'\boldsymbol{w} - 2\boldsymbol{x}_j'\boldsymbol{w}\alpha_j + [\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2\}$$

Completing the square in this expression with respect to α_j gives

$$-\frac{1}{2\sigma_e^2}\{c_j(\alpha_j - \hat{\alpha}_j)^2 + \boldsymbol{w}'\boldsymbol{w} - c_j\hat{\alpha}_j{}^2\}$$

where

$$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$

So,

$$f(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \propto \exp\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}}\}$$

Derive: full conditional for α_j

The exponential terms in the joint density can be written as:

$$-\frac{1}{2\sigma_e^2}\{\boldsymbol{w}'\boldsymbol{w} - 2\boldsymbol{x}_j'\boldsymbol{w}\alpha_j + [\boldsymbol{x}_j'\boldsymbol{x}_j + \frac{\sigma_e^2}{\sigma_\alpha^2}]\alpha_j^2\}$$

Completing the square in this expression with respect to α_j gives

$$-\frac{1}{2\sigma_e^2}\{c_j(\alpha_j - \hat{\alpha}_j)^2 + \boldsymbol{w}'\boldsymbol{w} - c_j\hat{\alpha}_j{}^2\}$$

where

$$\hat{\alpha}_j = \frac{\boldsymbol{x}_j'\boldsymbol{w}}{c_j}$$

So,

$$f(\alpha_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \sigma_e^2) \propto \exp\{-\frac{(\alpha_j - \hat{\alpha}_j)^2}{2\frac{\sigma_e^2}{c_j}}\}$$

Full conditional for σ_e^2

From Bayes' theorem,

$$f(\sigma_e^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}) = \frac{f(\sigma_e^2, \boldsymbol{y}, \mu, \boldsymbol{\alpha})}{f(\boldsymbol{y}, \mu, \boldsymbol{\alpha})}$$

$$\propto f(\boldsymbol{y}|\sigma_e^2, \mu, \boldsymbol{\alpha}) f(\sigma_e^2) f(\mu, \boldsymbol{\alpha})$$

where

$$f(\boldsymbol{y}|\sigma_e^2, \mu, \boldsymbol{\alpha}) \propto (\sigma_e^2)^{-n/2} \exp\{-\frac{(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)'(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}$$

and

$$f(\sigma_e^2) = \frac{(S_e^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)} (\sigma_e^2)^{-(2+\nu_e)/2} \exp(-\frac{\nu_e S_e^2}{2\sigma_e^2})$$

Full conditional for σ_e^2

From Bayes' theorem,

$$f(\sigma_e^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}) = \frac{f(\sigma_e^2, \boldsymbol{y}, \mu, \boldsymbol{\alpha})}{f(\boldsymbol{y}, \mu, \boldsymbol{\alpha})}$$

$$\propto f(\boldsymbol{y}|\sigma_e^2, \mu, \boldsymbol{\alpha}) f(\sigma_e^2) f(\mu, \boldsymbol{\alpha})$$

where

$$f(\boldsymbol{y}|\sigma_e^2, \mu, \boldsymbol{\alpha}) \propto (\sigma_e^2)^{-n/2} \exp\{-\frac{(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)'(\boldsymbol{w} - \boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}$$

and

$$f(\sigma_e^2) = \frac{(S_e^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)}(\sigma_e^2)^{-(2+\nu_e)/2} \exp(-\frac{\nu_e S_e^2}{2\sigma_e^2})$$

Full conditional for σ_e^2

From Bayes' theorem,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) = \frac{f(\sigma_e^2,\boldsymbol{y},\mu,\boldsymbol{\alpha})}{f(\boldsymbol{y},\mu,\boldsymbol{\alpha})}$$

$$\propto f(\boldsymbol{y}|\sigma_e^2,\mu,\boldsymbol{\alpha})f(\sigma_e^2)f(\mu,\boldsymbol{\alpha})$$

where

$$f(\boldsymbol{y}|\sigma_e^2,\mu,\boldsymbol{\alpha}) \propto (\sigma_e^2)^{-n/2}\exp\{-\frac{(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}$$

and

$$f(\sigma_e^2) = \frac{(S_e^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)}(\sigma_e^2)^{-(2+\nu_e)/2}\exp(-\frac{\nu_e S_e^2}{2\sigma_e^2})$$

Full conditional for σ_e^2

From Bayes' theorem,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) = \frac{f(\sigma_e^2,\boldsymbol{y},\mu,\boldsymbol{\alpha})}{f(\boldsymbol{y},\mu,\boldsymbol{\alpha})}$$

$$\propto f(\boldsymbol{y}|\sigma_e^2,\mu,\boldsymbol{\alpha})f(\sigma_e^2)f(\mu,\boldsymbol{\alpha})$$

where

$$f(\boldsymbol{y}|\sigma_e^2,\mu,\boldsymbol{\alpha}) \propto (\sigma_e^2)^{-n/2}\exp\{-\frac{(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)}{2\sigma_e^2}\}$$

and

$$f(\sigma_e^2) = \frac{(S_e^2\nu_e/2)^{\nu_e/2}}{\Gamma(\nu/2)}(\sigma_e^2)^{-(2+\nu_e)/2}\exp(-\frac{\nu_e S_e^2}{2\sigma_e^2})$$

Full conditional for σ_e^2

So,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \propto (\sigma_e^2)^{-(2+n+\nu_e)/2}\exp(-\frac{SSE+\nu_e S_e^2}{2\sigma_e^2})$$

where

$$SSE = (\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)$$

So,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \sim \tilde{\nu}_e\tilde{S}_e^2\chi_{\tilde{\nu}_e}^{-2}$$

where

$$\tilde{\nu}_e = n+\nu_e; \quad \tilde{S}_e^2 = \frac{SSE+\nu_e S_e^2}{\tilde{\nu}_e}$$

Full conditional for σ_e^2

So,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \propto (\sigma_e^2)^{-(2+n+\nu_e)/2}\exp(-\frac{SSE+\nu_e S_e^2}{2\sigma_e^2})$$

where

$$SSE = (\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)'(\boldsymbol{w}-\boldsymbol{x}_j\alpha_j)$$

So,

$$f(\sigma_e^2|\boldsymbol{y},\mu,\boldsymbol{\alpha}) \sim \tilde{\nu}_e\tilde{S}_e^2\chi_{\tilde{\nu}_e}^{-2}$$

where

$$\tilde{\nu}_e = n+\nu_e; \quad \tilde{S}_e^2 = \frac{SSE+\nu_e S_e^2}{\tilde{\nu}_e}$$

Alternative view of Normal prior

Consider fixed linear model:

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{e}$$

This can be also written as

$$\boldsymbol{y} = \begin{bmatrix} \mathbf{1} & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \boldsymbol{e}$$

Suppose we observe for each locus:

$$y_j^* = \alpha_j + \epsilon_j$$

Alternative view of Normal prior

Consider fixed linear model:

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{e}$$

This can be also written as

$$\boldsymbol{y} = \begin{bmatrix} \mathbf{1} & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \mu \\ \boldsymbol{\alpha} \end{bmatrix} + \boldsymbol{e}$$

Suppose we observe for each locus:

$$y_j^* = \alpha_j + \epsilon_j$$

Alternative view of Normal prior

Consider fixed linear model:

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{X}\boldsymbol{\alpha} + \boldsymbol{e}$$

This can be also written as

$$\boldsymbol{y} = \begin{bmatrix} \mathbf{1} & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \mu \\ \boldsymbol{\alpha} \end{bmatrix} + \boldsymbol{e}$$

Suppose we observe for each locus:

$$y_j^* = \alpha_j + \epsilon_j$$

Least Squares with Additional Data

Fixed linear model with the additional data:

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{y}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \begin{bmatrix} \boldsymbol{e} \\ \epsilon \end{bmatrix}$$

OLS Equations:

$$\begin{bmatrix} \boldsymbol{1}' & \boldsymbol{0}' \\ \boldsymbol{X}' & \boldsymbol{I}' \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n \frac{1}{\sigma_e^2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{1} & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}' & \boldsymbol{0}' \\ \boldsymbol{X}' & \boldsymbol{I}' \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n \frac{1}{\sigma_e^2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{y}^* \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{1}'\boldsymbol{1} & \boldsymbol{1}'\boldsymbol{X} \\ \boldsymbol{X}'\boldsymbol{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}'\boldsymbol{y} \\ \boldsymbol{X}'\boldsymbol{y} + \boldsymbol{y}^*\frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix}$$

Least Squares with Additional Data

Fixed linear model with the additional data:

$$\begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{y}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{1} & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \mu \\ \alpha \end{bmatrix} + \begin{bmatrix} \boldsymbol{e} \\ \epsilon \end{bmatrix}$$

OLS Equations:

$$\begin{bmatrix} \boldsymbol{1}' & \boldsymbol{0}' \\ \boldsymbol{X}' & \boldsymbol{I}' \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n \frac{1}{\sigma_e^2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{1} & \boldsymbol{X} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}' & \boldsymbol{0}' \\ \boldsymbol{X}' & \boldsymbol{I}' \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_n \frac{1}{\sigma_e^2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{y}^* \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{1}'\boldsymbol{1} & \boldsymbol{1}'\boldsymbol{X} \\ \boldsymbol{X}'\boldsymbol{1} & \boldsymbol{X}'\boldsymbol{X} + \boldsymbol{I}\frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} \boldsymbol{1}'\boldsymbol{y} \\ \boldsymbol{X}'\boldsymbol{y} + \boldsymbol{y}^* \frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix}$$

Least Squares with Additional Data

Fixed linear model with the additional data:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mu \\ \boldsymbol{\alpha} \end{bmatrix} + \begin{bmatrix} \mathbf{e} \\ \boldsymbol{\epsilon} \end{bmatrix}$$

OLS Equations:

$$\begin{bmatrix} \mathbf{1}' & \mathbf{0}' \\ \mathbf{X}' & \mathbf{I}' \end{bmatrix} \begin{bmatrix} \mathbf{I}_n \frac{1}{\sigma_e^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}' & \mathbf{0}' \\ \mathbf{X}' & \mathbf{I}' \end{bmatrix} \begin{bmatrix} \mathbf{I}_n \frac{1}{\sigma_e^2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k \frac{1}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{y}^* \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{X} \\ \mathbf{X}'\mathbf{1} & \mathbf{X}'\mathbf{X} + \mathbf{I}\frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{X}'\mathbf{y} + \mathbf{y}^* \frac{\sigma_e^2}{\sigma_\epsilon^2} \end{bmatrix}$$

Univariate-t

Prior:

$$(\alpha_j|\sigma_j^2) \sim \text{N}(0, \sigma_j^2)$$

$$\sigma_j^2 \sim \nu_\alpha S^2_{\nu_\alpha} \chi^{-2}_{\nu_\alpha}$$

Can show that the unconditional distribution for α_j is

$$\alpha_j \sim \text{(iid)}t(0, S^2_{\nu_\alpha}, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG pages 28,60)

This is Bayes-A (Meuwissen et al., 2001; Genetics 157:1819-1829)

Univariate-t

Prior:

$$(\alpha_j|\sigma_j^2) \sim \mathrm{N}(0, \sigma_j^2)$$

$$\sigma_j^2 \sim \nu_\alpha S_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}$$

Can show that the unconditional distribution for α_j is

$$\alpha_j \sim (\text{iid})t(0, S_{\nu_\alpha}^2, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG pages 28,60)

This is Bayes-A (Meuwissen et al., 2001; Genetics 157:1819-1829)

Univariate-t

Prior:

$$(\alpha_j|\sigma_j^2) \sim \text{N}(0, \sigma_j^2)$$

$$\sigma_j^2 \sim \nu_\alpha S_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}$$

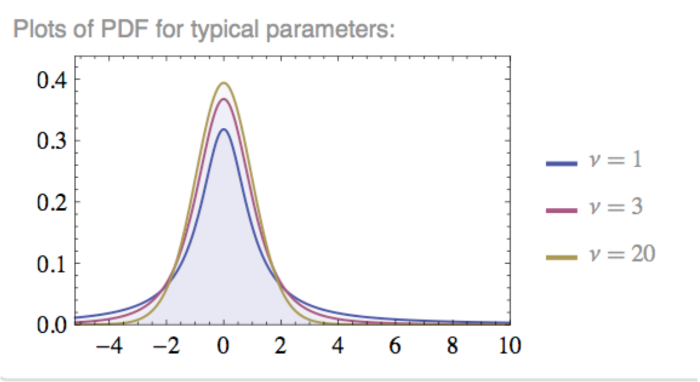
Can show that the unconditional distribution for α_j is

$$\alpha_j \sim (\text{iid})t(0, S_{\nu_\alpha}^2, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG pages 28,60)

This is Bayes-A (Meuwissen et al., 2001; Genetics 157:1819-1829)

Univariate-t



Generated by Wolfram|Alpha (www.wolframalpha.com)

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\xi = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \xi_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \xi, \sigma_e^2)$$

$$\propto f(\boldsymbol{y}|\mu, \boldsymbol{\alpha}, \xi, \sigma_e^2) f(\alpha_j|\sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \xi_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \xi_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \xi, \sigma_e^2)$$

$$\propto f(\boldsymbol{y} | \mu, \boldsymbol{\alpha}, \xi, \sigma_e^2) f(\alpha_j | \sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \xi_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2)$$

$$\propto f(\boldsymbol{y} | \mu, \boldsymbol{\alpha}, \xi, \sigma_e^2) f(\alpha_j | \sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \xi_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2)$$

$$\propto f(\boldsymbol{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2) f(\alpha_j | \sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2)$$

$$\propto f(\boldsymbol{y} | \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2) f(\alpha_j | \sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for single-site Gibbs

Full conditionals are the same as in the "Normal" model for μ, α_j, and σ_e^2. Let

$$\boldsymbol{\xi} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]$$

Full conditional conditional for σ_j^2:

$$f(\sigma_j^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}_{j\_}, \sigma_e^2) \propto f(\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2)$$

$$\propto f(\boldsymbol{y}|\mu, \boldsymbol{\alpha}, \boldsymbol{\xi}, \sigma_e^2) f(\alpha_j|\sigma_j^2) f(\sigma_j^2) f(\mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{j\_} \sigma_e^2)$$

$$\propto (\sigma_j^2)^{-1/2} \exp\{-\frac{\alpha_j^2}{2\sigma_j^2}\} (\sigma_j^2)^{-(2+\nu_\alpha)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

$$\propto (\sigma_j^2)^{-(2+\nu_\alpha+1)/2} \exp\{\frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{2\sigma_j^2}\}$$

Full conditional for σ_j^2

So,

$$(\sigma_j^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\xi}\_, \sigma_e^2) \sim \tilde{\nu}_\alpha \tilde{S}_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

where

$$\tilde{\nu}_\alpha = \nu_\alpha + 1$$

and

$$\tilde{S}_\alpha^2 = \frac{\alpha_j^2 + \nu_\alpha S_\alpha^2}{\tilde{\nu}_\alpha}$$

Multivariate-t

Prior:

$$(\alpha_j | \sigma^2_\alpha) \sim (\text{iid})\text{N}(0, \sigma^2_\alpha)$$

$$\sigma^2_\alpha \sim \nu_\alpha S^2_{\nu_\alpha} \chi^{-2}_{\nu_\alpha}$$

Can show that the unconditional distribution for α is

$$\boldsymbol{\alpha} \sim \text{multivariate-}t(\mathbf{0}, \boldsymbol{I} S^2_{\nu_\alpha}, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG page 60)

We will see later that this is Bayes-C with $\pi = 0$.

Multivariate-t

Prior:

$$(\alpha_j|\sigma^2_\alpha) \sim \text{(iid)N}(0, \sigma^2_\alpha)$$

$$\sigma^2_\alpha \sim \nu_\alpha S^2_{\nu_\alpha} \chi^{-2}_{\nu_\alpha}$$

Can show that the unconditional distribution for $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha} \sim \text{multivariate-}t(\mathbf{0}, \boldsymbol{I}S^2_{\nu_\alpha}, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG page 60)

We will see later that this is Bayes-C with $\pi = 0$.

Multivariate-t

Prior:

$$(\alpha_j|\sigma_\alpha^2) \sim \text{(iid)N}(0, \sigma_\alpha^2)$$

$$\sigma_\alpha^2 \sim \nu_\alpha S_{\nu_\alpha}^2 \chi_{\nu_\alpha}^{-2}$$

Can show that the unconditional distribution for $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha} \sim \text{multivariate-}t(\mathbf{0}, \boldsymbol{I}S_{\nu_\alpha}^2, \nu_\alpha)$$

(Sorensen and Gianola, 2002, LBMMQG page 60)

We will see later that this is Bayes-C with $\pi = 0$.

Full conditional for σ^2_α

We will see later that

$$(\sigma^2_\alpha | \boldsymbol{y}, \mu, \boldsymbol{\alpha}, \sigma^2_e) \sim \tilde{\nu}_\alpha \tilde{S}^2_\alpha \chi^{-2}_{\nu_\alpha}$$

where

$$\tilde{\nu}_\alpha = \nu_\alpha + k$$

and

$$\tilde{S}^2_\alpha = \frac{\boldsymbol{\alpha}'\boldsymbol{\alpha} + \nu_\alpha S^2_\alpha}{\tilde{\nu}_\alpha}$$

Spike and univariate-t

Prior:

$$(\alpha_j | \pi, \sigma_j^2) \begin{cases} \sim \mathrm{N}(0, \sigma_j^2) & \text{probability } (1-\pi), \\ = 0 & \text{probability } \pi \end{cases}$$

and

$$(\sigma_j^2 | \nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

Thus,

$$(\alpha_j | \pi)(\text{iid}) \begin{cases} \sim \text{univariate-}t(0, S_\alpha^2, \nu_\alpha) & \text{probability } (1-\pi), \\ = 0 & \text{probability } \pi \end{cases}$$

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)

Spike and univariate-t

Prior:

$$(\alpha_j|\pi, \sigma_j^2) \begin{cases} \sim \mathrm{N}(0, \sigma_j^2) & \text{probability } (1-\pi), \\ = 0 & \text{probability } \pi \end{cases}$$

and

$$(\sigma_j^2|\nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

Thus,

$$(\alpha_j|\pi)(\text{iid}) \begin{cases} \sim \text{univariate-}t(0, S_\alpha^2, \nu_\alpha) & \text{probability } (1-\pi), \\ = 0 & \text{probability } \pi \end{cases}$$

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)

Spike and univariate-t

Prior:

$$(\alpha_j|\pi,\sigma_j^2)\begin{cases}\sim \mathrm{N}(0,\sigma_j^2) & \text{probability } (1-\pi),\\ = 0 & \text{probability } \pi\end{cases}$$

and

$$(\sigma_j^2|\nu_\alpha, S_\alpha^2) \sim \nu_\alpha S_\alpha^2 \chi_{\nu_\alpha}^{-2}$$

Thus,

$$(\alpha_j|\pi)(\text{iid})\begin{cases}\sim \text{univariate-}t(0,S_\alpha^2,\nu_\alpha) & \text{probability } (1-\pi),\\ = 0 & \text{probability } \pi\end{cases}$$

This is Bayes-B (Meuwissen et al., 2001; Genetics 157:1819-1829)

Notation for sampling from mixture

The indicator variable δ_j is defined as

$$\delta_j = 1 \Rightarrow (\alpha_j|\sigma_j^2) \sim \mathrm{N}(0, \sigma_j^2)$$

and

$$\delta_j = 0 \Rightarrow (\alpha_j|\sigma_j^2) = 0$$

Sampling strategy in MHG (2001)

- Sampling σ_e^2 and μ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for σ_j^2 and α_j jointly from their full-conditional distribution.
- First, σ_j^2 is sampled from

$$f(\sigma_j^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{\_}, \sigma_e^2)$$

  using MH with prior as proposal.
- Then, α_j is sampled from its full-conditional, which is identical to that under the Normal prior

Sampling strategy in MHG (2001)

- Sampling σ_e^2 and μ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for σ_j^2 and α_j jointly from their full-conditional distribution.
- First, σ_j^2 is sampled from

$$f(\sigma_j^2|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{\_}, \sigma_e^2)$$

  using MH with prior as proposal.
- Then, α_j is sampled from its full-conditional, which is identical to that under the Normal prior

Sampling strategy in MHG (2001)

- Sampling σ_e^2 and μ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for σ_j^2 and α_j jointly from their full-conditional distribution.
- First, σ_j^2 is sampled from

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{\_}, \sigma_e^2)$$

  using MH with prior as proposal.
- Then, α_j is sampled from its full-conditional, which is identical to that under the Normal prior

Sampling strategy in MHG (2001)

- Sampling σ_e^2 and μ are as under the Normal prior.
- MHG proposed to use a Metropolis-Hastings sampler to draw samples for σ_j^2 and α_j jointly from their full-conditional distribution.
- First, σ_j^2 is sampled from

$$f(\sigma_j^2 | \boldsymbol{y}, \mu, \boldsymbol{\alpha}_{j\_}, \boldsymbol{\xi}_{\_}, \sigma_e^2)$$

  using MH with prior as proposal.
- Then, α_j is sampled from its full-conditional, which is identical to that under the Normal prior

MH acceptance probability when prior is used as proposal

Suppose we want to sample θ from $f(\theta|\boldsymbol{y})$ using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

$$\alpha = min(1, \frac{f(\theta_{can}|\boldsymbol{y})f(\theta^{t-1})}{f(\theta^{t-1}|\boldsymbol{y})f(\theta_{can})}$$

where $f(\theta)$ is the prior for θ. Using Bayes' theorem, the target density can be written as:

$$f(\theta|\boldsymbol{y}) \propto f(\boldsymbol{y}|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = min(1, \frac{f(\boldsymbol{y}|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(\boldsymbol{y}|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})}$$

MH acceptance probability when prior is used as proposal

Suppose we want to sample θ from $f(\theta|\boldsymbol{y})$ using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

$$\alpha = min(1, \frac{f(\theta_{can}|\boldsymbol{y})f(\theta^{t-1})}{f(\theta^{t-1}|\boldsymbol{y})f(\theta_{can})}$$

where $f(\theta)$ is the prior for θ. Using Bayes' theorem, the target density can be written as:

$$f(\theta|\boldsymbol{y}) \propto f(\boldsymbol{y}|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = min(1, \frac{f(\boldsymbol{y}|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(\boldsymbol{y}|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})}$$

MH acceptance probability when prior is used as proposal

Suppose we want to sample θ from $f(\theta|\boldsymbol{y})$ using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

$$\alpha = min(1, \frac{f(\theta_{can}|\boldsymbol{y})f(\theta^{t-1})}{f(\theta^{t-1}|\boldsymbol{y})f(\theta_{can})}$$

where $f(\theta)$ is the prior for θ. Using Bayes' theorem, the target density can be written as:

$$f(\theta|\boldsymbol{y}) \propto f(\boldsymbol{y}|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = min(1, \frac{f(\boldsymbol{y}|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(\boldsymbol{y}|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})}$$

MH acceptance probability when prior is used as proposal

Suppose we want to sample θ from $f(\theta|\boldsymbol{y})$ using the MH with its prior as proposal. Then, the MH acceptance probability becomes:

$$\alpha = min(1, \frac{f(\theta_{can}|\boldsymbol{y})f(\theta^{t-1})}{f(\theta^{t-1}|\boldsymbol{y})f(\theta_{can})}$$

where $f(\theta)$ is the prior for θ. Using Bayes' theorem, the target density can be written as:

$$f(\theta|\boldsymbol{y}) \propto f(\boldsymbol{y}|\theta)f(\theta)$$

Then, the acceptance probability becomes

$$\alpha = min(1, \frac{f(\boldsymbol{y}|\theta_{can})f(\theta_{can})f(\theta^{t-1})}{f(\boldsymbol{y}|\theta^{t-1})f(\theta^{t-1})f(\theta_{can})}$$

Sampling σ_j^2

Thus when the prior for σ_j^2 is used as the proposal, the MH acceptance probability becomes

$$\alpha = \min(1, \frac{f(\boldsymbol{y}|\sigma_{can}^2, \boldsymbol{\theta}_{j\_})}{f(\boldsymbol{y}|\sigma_j^2, \boldsymbol{\theta}_{j\_})})$$

where σ_{can}^2 is used to denote the candidate value for σ_j^2, and $\boldsymbol{\theta}_{j\_}$ all the other parameters. It can be shown that, α_j depends on $\boldsymbol{y}$ only through $r_j = \boldsymbol{x}_j'\boldsymbol{w}$ (page 30). Thus

$$f(\boldsymbol{y}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) \propto f(r_j|\sigma_j^2, \boldsymbol{\theta}_{j\_})$$

Sampling σ_j^2

Thus when the prior for σ_j^2 is used as the proposal, the MH acceptance probability becomes

$$\alpha = \min(1, \frac{f(\boldsymbol{y}|\sigma^2_{can}, \boldsymbol{\theta}_{j\_})}{f(\boldsymbol{y}|\sigma^2_j, \boldsymbol{\theta}_{j\_})})$$

where σ^2_{can} is used to denote the candidate value for σ_j^2, and $\boldsymbol{\theta}_{j\_}$ all the other parameters. It can be shown that, α_j depends on $\boldsymbol{y}$ only through $r_j = \boldsymbol{x}_j'\boldsymbol{w}$ (page 30). Thus

$$f(\boldsymbol{y}|\sigma^2_j, \boldsymbol{\theta}_{j\_}) \propto f(r_j|\sigma^2_j, \boldsymbol{\theta}_{j\_})$$

"Likelihood" for σ_j^2

Recall that

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'} = \boldsymbol{x}_j\alpha_j + \boldsymbol{e}$$

Then,

$$\mathrm{E}(\boldsymbol{w}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{0}$$

When $\delta = 1$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j\boldsymbol{x}_j'\sigma_j^2 + \boldsymbol{I}\sigma_e^2$$

and $\delta = 0$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{I}\sigma_e^2$$

"Likelihood" for σ_j^2

Recall that

$$\boldsymbol{w} = \boldsymbol{y} - \mathbf{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'} = \boldsymbol{x}_j\alpha_j + \boldsymbol{e}$$

Then,

$$\mathrm{E}(\boldsymbol{w}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = \mathbf{0}$$

When $\delta = 1$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j\boldsymbol{x}_j'\sigma_j^2 + \boldsymbol{I}\sigma_e^2$$

and $\delta = 0$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{I}\sigma_e^2$$

"Likelihood" for σ_j^2

Recall that

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'} = \boldsymbol{x}_j\alpha_j + \boldsymbol{e}$$

Then,

$$\mathrm{E}(\boldsymbol{w}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{0}$$

When $\delta = 1$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j\boldsymbol{x}_j'\sigma_j^2 + \boldsymbol{I}\sigma_e^2$$

and $\delta = 0$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{I}\sigma_e^2$$

"Likelihood" for σ_j^2

Recall that

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'} = \boldsymbol{x}_j\alpha_j + \boldsymbol{e}$$

Then,

$$\mathrm{E}(\boldsymbol{w}|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{0}$$

When $\delta = 1$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j\boldsymbol{x}_j'\sigma_j^2 + \boldsymbol{I}\sigma_e^2$$

and $\delta = 0$:

$$\mathrm{Var}(\boldsymbol{w}|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{I}\sigma_e^2$$

"Likelihood" for σ_j^2

So,

$$\mathrm{E}(r_j|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = 0$$

and

$$\mathrm{Var}(r_j|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = (\boldsymbol{x}_j'\boldsymbol{x}_j)^2\sigma_j^2 + \boldsymbol{x}_j'\boldsymbol{x}_j\sigma_e^2 = v_1$$

$$\mathrm{Var}(r_j|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j'\boldsymbol{x}_j\sigma_e^2 = v_0$$

So,

$$f(r_j|\delta_j, \sigma_j^2, \boldsymbol{\theta}_{j\_}) \propto (v_\delta)^{-1/2}\exp\{-\frac{r_j^2}{2v_\delta}\}$$

"Likelihood" for σ_j^2

So,

$$E(r_j|\sigma_j^2, \boldsymbol{\theta}_{j\_}) = 0$$

and

$$\text{Var}(r_j|\delta_j = 1, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = (\boldsymbol{x}_j'\boldsymbol{x}_j)^2\sigma_j^2 + \boldsymbol{x}_j'\boldsymbol{x}_j\sigma_e^2 = v_1$$

$$\text{Var}(r_j|\delta_j = 0, \sigma_j^2, \boldsymbol{\theta}_{j\_}) = \boldsymbol{x}_j'\boldsymbol{x}_j\sigma_e^2 = v_0$$

So,

$$f(r_j|\delta_j, \sigma_j^2, \boldsymbol{\theta}_{j\_}) \propto (v_\delta)^{-1/2}\exp\{-\frac{r_j^2}{2v_\delta}\}$$

Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.

Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.

Alternative View of Prior in BayesB

- How much information is being added by the prior?
- BayesB is identical to ML with additional data!
- Can “see” how much additional data in BayesB prior.

Maximum Likelihood with Additional Data

- Suppose at locus j, $\delta_j = 1$, and we observe additional data:

$$\boldsymbol{u}_j \sim N(\boldsymbol{0}, \boldsymbol{I}_q\sigma_j^2)$$

- Assume that only unknown is σ_j^2
- So, adjust phenotypes as:

$$\boldsymbol{w} = \boldsymbol{y} - \boldsymbol{1}\mu - \sum_{j' \neq j} \boldsymbol{x}_{j'}\alpha_{j'}$$

- Likelihood:

$$L(\sigma_j^2; \boldsymbol{w}, \boldsymbol{u}_j) = L(\sigma_j^2; \hat{\alpha}_j, \boldsymbol{u}_j)$$

Likelihood with Additional Data

- $$L(\sigma_j^2; \hat{\alpha}_j, \boldsymbol{u}_j) \propto f_1(\hat{\alpha}_j|\sigma_j^2) \times f_2(\boldsymbol{u}_j|\sigma_j^2)$$

- $$f_2(\boldsymbol{u}_j|\sigma_j^2) \propto (\sigma_j^2)^{-q/2} \exp[\frac{-\boldsymbol{u}_j'\boldsymbol{u}_j}{2\sigma_j^2}]$$
  $$\propto (\sigma_j^2)^{-[\nu/2+1]} \exp[\frac{-\nu S^2}{2\sigma_j^2}]$$

- $\nu = q - 2, S^2 = \frac{\boldsymbol{u}_j'\boldsymbol{u}_j}{\nu}$

Likelihood with Additional Data

- $$L(\sigma_j^2; \hat{\alpha}_j, \boldsymbol{u}_j) \propto f_1(\hat{\alpha}_j|\sigma_j^2) \times f_2(\boldsymbol{u}_j|\sigma_j^2)$$
- $$\begin{aligned} f_2(\boldsymbol{u}_j|\sigma_j^2) &\propto (\sigma_j^2)^{-q/2} \exp[\frac{-\boldsymbol{u}_j'\boldsymbol{u}_j}{2\sigma_j^2}] \\ &\propto (\sigma_j^2)^{-[\nu/2+1]} \exp[\frac{-\nu S^2}{2\sigma_j^2}] \end{aligned}$$
- $\nu = q - 2, S^2 = \frac{\boldsymbol{u}_j'\boldsymbol{u}_j}{\nu}$

Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling σ_j^2, we

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample σ_j^2 from a scaled inverse chi-squared distribution with
  - scale parameter = $\sigma_j^{2(t-1)}/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter = S_α^2 and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$

Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling σ_j^2, we

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample σ_j^2 from a scaled inverse chi-squared distribution with
  - scale parameter = $\sigma_j^{2(t-1)}/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter = S_α^2 and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$

Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling σ_j^2, we

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample σ_j^2 from a scaled inverse chi-squared distribution with
  - scale parameter = $\sigma_j^{2(t-1)}/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter = S_α^2 and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$

Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling σ_j^2, we

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample σ_j^2 from a scaled inverse chi-squared distribution with
  - scale parameter = $\sigma_j^{2(t-1)}/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter = S_α^2 and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$

Alternative algorithm for spike and univariate-t

Rather than use the prior as the proposal for sampling σ_j^2, we

- sample $\delta_j = 1$ with probability 0.5
- when $\delta = 1$, sample σ_j^2 from a scaled inverse chi-squared distribution with
  - scale parameter = $\sigma_j^{2(t-1)}/2$ and 4 degrees of freedom when $\delta_j^{(t-1)} = 1$, and
  - scale parameter = S_α^2 and 4 degrees of freedom when $\delta_j^{(t-1)} = 0$

Multivariate-t mixture

Prior:

$$(\alpha_j|\pi,\sigma_\alpha^2)\begin{cases}\sim \mathrm{N}(0,\sigma_\alpha^2) & \text{probability } (1-\pi),\\ = 0 & \text{probability } \pi\end{cases}$$

and

$$(\sigma_\alpha^2|\nu_\alpha, S_\alpha^2)\sim \nu_\alpha S_\alpha^2\chi_{\nu_\alpha}^{-2}$$

Further,

$$\pi \sim \mathrm{Uniform}(0,1)$$

- The α_j variables with their corresponding $\delta_j = 1$ will follow a multivariate-t distribution.
- This is what we have called Bayes-Cπ

Multivariate-t mixture

Prior:

$$(\alpha_j|\pi, \sigma^2_\alpha) \begin{cases} \sim \mathrm{N}(0, \sigma^2_\alpha) & \text{probability } (1-\pi), \\ = 0 & \text{probability } \pi \end{cases}$$

and

$$(\sigma^2_\alpha|\nu_\alpha, S^2_\alpha) \sim \nu_\alpha S^2_\alpha \chi^{-2}_{\nu_\alpha}$$

Further,

$$\pi \sim \mathrm{Uniform}(0,1)$$

- The α_j variables with their corresponding $\delta_j = 1$ will follow a multivariate-t distribution.
- This is what we have called Bayes-Cπ

Multivariate-t mixture

Prior:

$$(\alpha_j|\pi,\sigma_\alpha^2)\begin{cases}\sim \mathrm{N}(0,\sigma_\alpha^2) & \text{probability } (1-\pi),\\ = 0 & \text{probability } \pi\end{cases}$$

and

$$(\sigma_\alpha^2|\nu_\alpha,S_\alpha^2)\sim \nu_\alpha S_\alpha^2\chi_{\nu_\alpha}^{-2}$$

Further,

$$\pi \sim \mathrm{Uniform}(0,1)$$

- The α_j variables with their corresponding $\delta_j = 1$ will follow a multivariate-t distribution.
- This is what we have called Bayes-Cπ

Multivariate-t mixture

Prior:

$$(\alpha_j|\pi,\sigma^2_\alpha)\begin{cases}\sim \mathrm{N}(0,\sigma^2_\alpha) & \text{probability } (1-\pi),\\ = 0 & \text{probability } \pi\end{cases}$$

and

$$(\sigma^2_\alpha|\nu_\alpha, S^2_\alpha)\sim \nu_\alpha S^2_\alpha \chi^{-2}_{\nu_\alpha}$$

Further,

$$\pi \sim \mathrm{Uniform}(0,1)$$

- The α_j variables with their corresponding $\delta_j = 1$ will follow a multivariate-t distribution.
- This is what we have called Bayes-Cπ

Full conditionals for single-site Gibbs

Full-conditional distributions for μ, $\boldsymbol{\alpha}$, and σ_e^2 are as with the Normal prior.

Full-conditional for δ_j:

$$\Pr(\delta_j|\mathbf{y},\mu,\boldsymbol{\alpha}_{-j},\boldsymbol{\delta}_{-j},\sigma_\alpha^2,\sigma_e^2,\pi) = \Pr(\delta_j|r_j,\boldsymbol{\theta}_{j\_})$$

$$\Pr(\delta_j|r_j,\boldsymbol{\theta}_{j\_}) = \frac{f(\delta_j, r_j|\boldsymbol{\theta}_{j\_})}{f(r_j|\boldsymbol{\theta}_{j\_})}$$

$$= \frac{f(r_j|\delta_j,\boldsymbol{\theta}_{j\_})\Pr(\delta_j|\pi)}{f(r_j|\delta_j=0,\boldsymbol{\theta}_{j\_})\pi + f(r_j|\delta_j=1,\boldsymbol{\theta}_{j\_})(1-\pi)}$$

Full conditionals for single-site Gibbs

Full-conditional distributions for μ, $\boldsymbol{\alpha}$, and σ_e^2 are as with the Normal prior.
Full-conditional for δ_j:

$$\Pr(\delta_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{-j}, \boldsymbol{\delta}_{-j}, \sigma_\alpha^2, \sigma_e^2, \pi) = \Pr(\delta_j|r_j, \boldsymbol{\theta}_{j\_})$$

$$\Pr(\delta_j|r_j, \boldsymbol{\theta}_{j\_}) = \frac{f(\delta_j, r_j|\boldsymbol{\theta}_{j\_})}{f(r_j|\boldsymbol{\theta}_{j\_})}$$

$$= \frac{f(r_j|\delta_j, \boldsymbol{\theta}_{j\_})\Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \boldsymbol{\theta}_{j\_})\pi + f(r_j|\delta_j = 1, \boldsymbol{\theta}_{j\_})(1-\pi)}$$

Full conditionals for single-site Gibbs

Full-conditional distributions for μ, $\boldsymbol{\alpha}$, and σ_e^2 are as with the Normal prior.
Full-conditional for δ_j:

$$\Pr(\delta_j | \boldsymbol{y}, \mu, \boldsymbol{\alpha}_{-j}, \boldsymbol{\delta}_{-j}, \sigma_\alpha^2, \sigma_e^2, \pi) = \Pr(\delta_j | r_j, \boldsymbol{\theta}_{j\_})$$

$$\Pr(\delta_j | r_j, \boldsymbol{\theta}_{j\_}) = \frac{f(\delta_j, r_j | \boldsymbol{\theta}_{j\_})}{f(r_j | \boldsymbol{\theta}_{j\_})}$$

$$= \frac{f(r_j | \delta_j, \boldsymbol{\theta}_{j\_}) \Pr(\delta_j | \pi)}{f(r_j | \delta_j = 0, \boldsymbol{\theta}_{j\_})\pi + f(r_j | \delta_j = 1, \boldsymbol{\theta}_{j\_})(1 - \pi)}$$

Full conditionals for single-site Gibbs

Full-conditional distributions for μ, $\boldsymbol{\alpha}$, and σ_e^2 are as with the Normal prior.
Full-conditional for δ_j:

$$\Pr(\delta_j|\boldsymbol{y}, \mu, \boldsymbol{\alpha}_{-j}, \boldsymbol{\delta}_{-j}, \sigma_\alpha^2, \sigma_e^2, \pi) = \Pr(\delta_j|r_j, \boldsymbol{\theta}_{j\_})$$

$$\Pr(\delta_j|r_j, \boldsymbol{\theta}_{j\_}) = \frac{f(\delta_j, r_j|\boldsymbol{\theta}_{j\_})}{f(r_j|\boldsymbol{\theta}_{j\_})}$$

$$= \frac{f(r_j|\delta_j, \boldsymbol{\theta}_{j\_}) \Pr(\delta_j|\pi)}{f(r_j|\delta_j = 0, \boldsymbol{\theta}_{j\_})\pi + f(r_j|\delta_j = 1, \boldsymbol{\theta}_{j\_})(1-\pi)}$$

Full conditional for σ^2_α

This can be written as

$$f(\sigma^2_\alpha|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\boldsymbol{y}|\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) f(\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

But, can see that

$$f(\boldsymbol{y}|\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\boldsymbol{y}|\mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

So,

$$f(\sigma^2_\alpha|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

Note that σ^2_α appears only in $f(\boldsymbol{\alpha}|\sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

$$f(\boldsymbol{\alpha}|\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-k/2} \exp\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma^2_\alpha}\}$$

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\}$$

Full conditional forσ_α^2

This can be written as

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

But, can see that

$$f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

So,

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

Note that σ_α^2 appears only in $f(\boldsymbol{\alpha}|\sigma_\alpha^2)$ and $f(\sigma_\alpha^2)$:

$$f(\boldsymbol{\alpha}|\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-k/2}\exp\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma_\alpha^2}\}$$

and

$$f(\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-(\nu_\alpha+2)/2}\exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\}$$

Full conditional for σ_α^2

This can be written as

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

But, can see that

$$f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

So,

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

Note that σ_α^2 appears only in $f(\alpha|\sigma_\alpha^2)$ and $f(\sigma_\alpha^2)$:

$$f(\boldsymbol{\alpha}|\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-k/2}\exp\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma_\alpha^2}\}$$

and

$$f(\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-(\nu_\alpha+2)/2}\exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\}$$

Full conditional for σ_α^2

This can be written as

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

But, can see that

$$f(\boldsymbol{y}|\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\boldsymbol{y}|\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

So,

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto f(\sigma_\alpha^2,\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2)$$

Note that σ_α^2 appears only in $f(\boldsymbol{\alpha}|\sigma_\alpha^2)$ and $f(\sigma_\alpha^2)$:

$$f(\boldsymbol{\alpha}|\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-k/2}\exp\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma_\alpha^2}\}$$

and

$$f(\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-(\nu_\alpha+2)/2}\exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\}$$

Full conditional forσ^2_α

This can be written as

$$f(\sigma^2_\alpha|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\boldsymbol{y}|\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)f(\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

But, can see that

$$f(\boldsymbol{y}|\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\boldsymbol{y}|\mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

So,

$$f(\sigma^2_\alpha|\boldsymbol{y}, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e) \propto f(\sigma^2_\alpha, \mu, \boldsymbol{\alpha}, \boldsymbol{\delta}, \sigma^2_e)$$

Note that σ^2_α appears only in $f(\boldsymbol{\alpha}|\sigma^2_\alpha)$ and $f(\sigma^2_\alpha)$:

$$f(\boldsymbol{\alpha}|\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-k/2} \exp\{-\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}}{2\sigma^2_\alpha}\}$$

and

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\}$$

Full conditional for σ_α^2

Combining these two densities gives:

$$f(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \propto (\sigma_\alpha^2)^{-(k+\nu_\alpha+2)/2}\exp\{\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}+\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\}$$

So,

$$(\sigma_\alpha^2|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma_e^2) \sim \tilde{\nu}_\alpha \tilde{S}_\alpha^2 \chi_{\tilde{\nu}_\alpha}^{-2}$$

where

$$\tilde{\nu}_\alpha = k + \nu_\alpha$$

and

$$\tilde{S}_\alpha^2 = \frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}+\nu_\alpha S_\alpha^2}{\tilde{\nu}_\alpha}$$

Full conditional for σ^2_α

Combining these two densities gives:

$$f(\sigma^2_\alpha|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma^2_e) \propto (\sigma^2_\alpha)^{-(k+\nu_\alpha+2)/2}\exp\{\frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}+\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\}$$

So,

$$(\sigma^2_\alpha|\boldsymbol{y},\mu,\boldsymbol{\alpha},\boldsymbol{\delta},\sigma^2_e) \sim \tilde{\nu}_\alpha\tilde{S}^2_\alpha\chi^{-2}_{\tilde{\nu}_\alpha}$$

where

$$\tilde{\nu}_\alpha = k+\nu_\alpha$$

and

$$\tilde{S}^2_\alpha = \frac{\boldsymbol{\alpha}'\boldsymbol{\alpha}+\nu_\alpha S^2_\alpha}{\tilde{\nu}_\alpha}$$

Hyper parameter: S_α^2

If σ^2 is distributed as a scaled, inverse chi-square random variable with scale parameter S^2 and degrees of freedom ν

$$E(\sigma^2) = \frac{\nu S^2}{\nu - 2}$$

Recall that under some assumptions

$$\sigma_\alpha^2 = \frac{V_a}{\sum_j 2p_jq_j}$$

So, we take

$$S_\alpha^2 = \frac{(\nu_\alpha - 2)V_a}{\nu_\alpha k(1-\pi)2\overline{pq}}$$

Hyper parameter: S_α^2

If σ^2 is distributed as a scaled, inverse chi-square random variable with scale parameter S^2 and degrees of freedom ν

$$\mathrm{E}(\sigma^2) = \frac{\nu S^2}{\nu - 2}$$

Recall that under some assumptions

$$\sigma_\alpha^2 = \frac{V_a}{\sum_j 2p_jq_j}$$

So, we take

$$S_\alpha^2 = \frac{(\nu_\alpha - 2)V_a}{\nu_\alpha k(1-\pi)2\overline{pq}}$$

Hyper parameter: S_α^2

If σ^2 is distributed as a scaled, inverse chi-square random variable with scale parameter S^2 and degrees of freedom ν

$$\mathrm{E}(\sigma^2) = \frac{\nu S^2}{\nu - 2}$$

Recall that under some assumptions

$$\sigma_\alpha^2 = \frac{V_a}{\sum_j 2p_j q_j}$$

So, we take

$$S_\alpha^2 = \frac{(\nu_\alpha - 2)V_a}{\nu_\alpha k(1-\pi)2\overline{pq}}$$

Full conditional for π

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) \propto f(\boldsymbol{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e) f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e)$$

But,

- Conditional on δ the likelihood is free of π
- Further, π only appears in probability of the vector of bernoulli variables: δ

Thus,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) = \pi^{(k-m)}(1-\pi)^m$$

where $m = \delta'\delta$, and k is the number of markers. Thus, π is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$.

Full conditional for π

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2, \boldsymbol{y}) \propto f(\boldsymbol{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2) f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2)$$

But,

- Conditional on δ the likelihood is free of π
- Further, π only appears in probability of the vector of bernoulli variables: δ

Thus,

$$f(\pi|\delta, \mu, \alpha, \sigma_\alpha^2, \sigma_e^2, \boldsymbol{y}) = \pi^{(k-m)}(1-\pi)^m$$

where $m = \delta'\delta$, and k is the number of markers. Thus, π is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$.

Full conditional for π

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) \propto f(\boldsymbol{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e) f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e)$$

But,

- Conditional on δ the likelihood is free of π
- Further, π only appears in probability of the vector of bernoulli variables: $\boldsymbol{\delta}$

Thus,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) = \pi^{(k-m)}(1-\pi)^m$$

where $m = \delta'\delta$, and k is the number of markers. Thus, π is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$.

Full conditional for π

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) \propto f(\boldsymbol{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e) f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e)$$

But,

- Conditional on δ the likelihood is free of π
- Further, π only appears in probability of the vector of bernoulli variables: $\boldsymbol{\delta}$

Thus,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma^2_\alpha, \sigma^2_e, \boldsymbol{y}) = \pi^{(k-m)}(1-\pi)^m$$

where $m = \delta'\delta$, and k is the number of markers. Thus, π is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$.

Full conditional for π

Using Bayes' theorem,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2, \boldsymbol{y}) \propto f(\boldsymbol{y}|\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2) f(\pi, \boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2)$$

But,

- Conditional on δ the likelihood is free of π
- Further, π only appears in probability of the vector of bernoulli variables: $\boldsymbol{\delta}$

Thus,

$$f(\pi|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2, \boldsymbol{y}) = \pi^{(k-m)}(1-\pi)^m$$

where $m = \boldsymbol{\delta}'\boldsymbol{\delta}$, and k is the number of markers. Thus, π is sampled from a beta distribution with $a = k - m + 1$ and $b = m + 1$.

BayesCπ with Unknown S_α^2

- Prior for S_α^2: Gamma(a,b)

$$f(S_\alpha^2|a,b) \propto b^a (S_\alpha^2)^{a-1} \exp\{-bS_\alpha^2\}$$

- Using Bayes theorem,

$$f(S_\alpha^2|\boldsymbol{\delta}, \mu, \boldsymbol{\alpha}, \sigma_\alpha^2, \sigma_e^2, \boldsymbol{y}) \propto f(\boldsymbol{y}|S_\alpha^2, \sigma_\alpha^2, \ldots) f(S_\alpha^2, \sigma^2 \ldots)$$

- Given $\mu, \boldsymbol{\alpha}$, and σ_e^2, $f(\boldsymbol{y}|S_\alpha^2, \sigma_\alpha^2, \ldots)$ does not depend on S_α^2.
- In $f(S_\alpha^2, \sigma^2 \ldots)$, S_α^2 is only in $f(S_\alpha^2|a,b)$ and $f(\sigma_\alpha^2|S_\alpha^2, \nu_\alpha)$

BayesCπ with Unknown S_α^2

- Prior for S_α^2: Gamma(a,b)

$$f(S_\alpha^2|a,b) \propto b^a (S_\alpha^2)^{a-1} \exp\{-bS_\alpha^2\}$$

- Using Bayes theorem,

$$f(S_\alpha^2|\boldsymbol{\delta},\mu,\boldsymbol{\alpha},\sigma_\alpha^2,\sigma_e^2,\boldsymbol{y}) \propto f(\boldsymbol{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)f(S_\alpha^2,\sigma^2\ldots)$$

- Given μ,α, and σ_e^2, $f(\boldsymbol{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)$ does not depend on S_α^2.
- In $f(S_\alpha^2,\sigma^2\ldots)$, S_α^2 is only in $f(S_\alpha^2|a,b)$ and $f(\sigma_\alpha^2|S_\alpha^2,\nu_\alpha)$

BayesCπ with Unknown S_α^2

- Prior for S_α^2: Gamma(a,b)

$$f(S_\alpha^2|a,b) \propto b^a (S_\alpha^2)^{a-1} \exp\{-bS_\alpha^2\}$$

- Using Bayes theorem,

$$f(S_\alpha^2|\boldsymbol{\delta},\mu,\boldsymbol{\alpha},\sigma_\alpha^2,\sigma_e^2,\mathbf{y}) \propto f(\mathbf{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)f(S_\alpha^2,\sigma^2\ldots)$$

- Given $\mu,\boldsymbol{\alpha}$, and σ_e^2, $f(\mathbf{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)$ does not depend on S_α^2.
- In $f(S_\alpha^2,\sigma^2\ldots)$, S_α^2 is only in $f(S_\alpha^2|a,b)$ and $f(\sigma_\alpha^2|S_\alpha^2,\nu_\alpha)$

BayesCπ with Unknown S_α^2

- Prior for S_α^2: Gamma(a,b)

$$f(S_\alpha^2|a,b) \propto b^a (S_\alpha^2)^{a-1} \exp\{-bS_\alpha^2\}$$

- Using Bayes theorem,

$$f(S_\alpha^2|\boldsymbol{\delta},\mu,\boldsymbol{\alpha},\sigma_\alpha^2,\sigma_e^2,\boldsymbol{y}) \propto f(\boldsymbol{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)f(S_\alpha^2,\sigma^2\ldots)$$

- Given $\mu,\boldsymbol{\alpha}$, and σ_e^2, $f(\boldsymbol{y}|S_\alpha^2,\sigma_\alpha^2,\ldots)$ does not depend on S_α^2.
- In $f(S_\alpha^2,\sigma^2\ldots)$, S_α^2 is only in $f(S_\alpha^2|a,b)$ and $f(\sigma_\alpha^2|S_\alpha^2,\nu_\alpha)$

BayesCπ with Unknown S^2_α

- Prior for S^2_α: Gamma(a,b)

$$f(S^2_\alpha | a, b) \propto b^a (S^2_\alpha)^{a-1} \exp\{-bS^2_\alpha\}$$

- Prior for σ^2_α:

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\}$$

- Combining these gives:

$$f(S^2_\alpha | \sigma^2_\alpha, \boldsymbol{y}, \ldots) \propto S^{2(a-1+\nu/2)}_\alpha \exp\{-S^2_\alpha(\frac{\nu_\alpha}{2\sigma^2_\alpha} + b)\}$$

BayesCπ with Unknown S^2_α

- Prior for S^2_α: Gamma(a,b)

$$f(S^2_\alpha | a, b) \propto b^a (S^2_\alpha)^{a-1} \exp\{-bS^2_\alpha\}$$

- Prior for σ^2_α:

$$f(\sigma^2_\alpha) \propto (\sigma^2_\alpha)^{-(\nu_\alpha+2)/2} \exp\{\frac{\nu_\alpha S^2_\alpha}{2\sigma^2_\alpha}\}$$

- Combining these gives:

$$f(S^2_\alpha | \sigma^2_\alpha, \mathbf{y}, \ldots) \propto S_\alpha^{2(a-1+\nu/2)} \exp\{-S^2_\alpha(\frac{\nu_\alpha}{2\sigma^2_\alpha} + b)\}$$

BayesCπ with Unknown S_α^2

- Prior for S_α^2: Gamma(a,b)

$$f(S_\alpha^2|a,b) \propto b^a (S_\alpha^2)^{a-1} \exp\{-bS_\alpha^2\}$$

- Prior for σ_α^2:

$$f(\sigma_\alpha^2) \propto (\sigma_\alpha^2)^{-(\nu_\alpha+2)/2} \exp\{\frac{\nu_\alpha S_\alpha^2}{2\sigma_\alpha^2}\}$$

- Combining these gives:

$$f(S_\alpha^2|\sigma_\alpha^2, \boldsymbol{y}, \ldots) \propto S_\alpha^{2(a-1+\nu/2)} \exp\{-S_\alpha^2(\frac{\nu_\alpha}{2\sigma_\alpha^2} + b)\}$$

BayesCπ with Unknown S^2_α

So, $f(S^2_\alpha | a, b)$ is Gamma(a*,b*), where

$$a* = a + \nu_\alpha / 2$$

and

$$b* = b + \frac{\nu_\alpha}{2\sigma^2_\alpha}$$

Simulation I

- 2000 unlinked loci in LE
- 10 of these are QTL: $\pi = 0.995$
- $h^2 = 0.5$
- Locus effects estimated from 250 individuals

Simulation I

- 2000 unlinked loci in LE
- 10 of these are QTL: $\pi = 0.995$
- $h^2 = 0.5$
- Locus effects estimated from 250 individuals

Simulation I

- 2000 unlinked loci in LE
- 10 of these are QTL: $\pi = 0.995$
- $h^2 = 0.5$
- Locus effects estimated from 250 individuals

Simulation I

- 2000 unlinked loci in LE
- 10 of these are QTL: $\pi = 0.995$
- $h^2 = 0.5$
- Locus effects estimated from 250 individuals

Results for Bayes-B

Correlations between true and predicted additive genotypic values estimated from 32 replications

π	S^2	Correlation
0.995	0.2	0.91 (0.009)
0.8	0.2	0.86 (0.009)
0.0	0.2	0.80 (0.013)
0.995	2.0	0.90 (0.007)
0.8	2.0	0.77 (0.009)
0.0	2.0	0.35 (0.022)

Simulation II

- 2000 unlinked loci with Q loci having effect on trait
- N is the size of training data set
- Heritability = 0.5
- Validation in an independent data set with 1000 individuals
- Bayes-B and Bayes-Cπ with $\pi = 0.5$

Simulation II

- 2000 unlinked loci with *Q* loci having effect on trait
- *N* is the size of training data set
- Heritability = 0.5
- Validation in an independent data set with 1000 individuals
- Bayes-B and Bayes-Cπ with $\pi = 0.5$

Simulation II

- 2000 unlinked loci with Q loci having effect on trait
- N is the size of training data set
- Heritability = 0.5
- Validation in an independent data set with 1000 individuals
- Bayes-B and Bayes-Cπ with $\pi = 0.5$

Simulation II

- 2000 unlinked loci with Q loci having effect on trait
- N is the size of training data set
- Heritability = 0.5
- Validation in an independent data set with 1000 individuals
- Bayes-B and Bayes-Cπ with $\pi = 0.5$

Simulation II

- 2000 unlinked loci with Q loci having effect on trait
- N is the size of training data set
- Heritability = 0.5
- Validation in an independent data set with 1000 individuals
- Bayes-B and Bayes-Cπ with $\pi = 0.5$

Results

Results from 15 replications

				Corr($g, \hat{g}$)	
N	Q	π	$\hat{\pi}$	Bayes-Cπ	Bayes-B
2000	10	0.995	0.994	0.995	0.937
2000	200	0.90	0.899	0.866	0.834
2000	1900	0.05	0.202	0.613	0.571
4000	1900	0.05	0.096	0.763	0.722

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ^2_α)
  - $\sigma^2_\alpha = \frac{\sigma^2_g}{50 2\bar{p}\bar{q}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ_α^2)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{502\bar{p}\bar{q}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ_α^2)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{502\bar{pq}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ^2_α)
  - $\sigma^2_\alpha = \frac{\sigma^2_g}{50 2\bar{p}\bar{q}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ_α^2)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{502\bar{pq}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ_α^2)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{50 2\bar{p}\bar{q}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Simulation III

- Genotypes: 50k SNPs from 1086 Purebred Angus animals, ISU
- Phenotypes:
  - QTL simulated from 50 randomly sampled SNPs
  - substitution effect sampled from N(0,σ_α^2)
  - $\sigma_\alpha^2 = \frac{\sigma_g^2}{502\bar{pq}}$
  - $h^2 = 0.25$
- QTL were included in the marker panel
- Marker effects were estimated for 50k SNPs

Validation

- Genotypes: 50k SNPs from 984 crossbred animals, CMP
- Additive genetic merit (g_i) computed from the 50 QTL
- Additive genetic merit predicted ($\hat{g}_i$) using estimated effects for 50k SNP panel

Validation

- Genotypes: 50k SNPs from 984 crossbred animals, CMP
- Additive genetic merit (g_i) computed from the 50 QTL
- Additive genetic merit predicted ($\hat{g}_i$) using estimated effects for 50k SNP panel

Validation

- Genotypes: 50k SNPs from 984 crossbred animals, CMP
- Additive genetic merit (g_i) computed from the 50 QTL
- Additive genetic merit predicted ($\hat{g}_i$) using estimated effects for 50k SNP panel

Results

Correlations between g_i and $\hat{g}_i$ estimated from 3 replications

	Correlation	
π	Bayes-B	Bayes-C
0.999	0.86	0.86
0.25	0.70	0.26

BayesCπ:

- $\hat{\pi} = 0.999$
- Correlation = 0.86

Results

Correlations between g_i and $\hat{g}_i$ estimated from 3 replications

	Correlation	
π	Bayes-B	Bayes-C
0.999	0.86	0.86
0.25	0.70	0.26

BayesCπ:

- $\hat{\pi} = 0.999$
- Correlation = 0.86

Results

Correlations between g_i and $\hat{g}_i$ estimated from 3 replications

	Correlation	
π	Bayes-B	Bayes-C
0.999	0.86	0.86
0.25	0.70	0.26

BayesCπ:

- $\hat{\pi} = 0.999$
- Correlation = 0.86

Results

Correlations between g_i and $\hat{g}_i$ estimated from 3 replications

	Correlation	
π	Bayes-B	Bayes-C
0.999	0.86	0.86
0.25	0.70	0.26

BayesCπ:

- $\hat{\pi} = 0.999$
- Correlation = 0.86