

METROPOLIS-HASTING ALGORITHM

PROBLEM DESCRIPTION

4 장을 통해 배운 Metropolis-Hasting 알고리즘을 활용하여 다음과 같은 문제를 푼다.

$$X_1, \dots, X_{10} \sim i.i.d. N(\mu, \sigma^2)$$

$$\text{사전분포: } \mu \sim N(\mu_0, \sigma_0^2), \sigma^2 \sim IG(a, b)$$

$$\text{관측치: } (x_1, \dots, x_{10}) = (10, 13, 15, 11, 9, 18, 20, 17, 23, 31)$$

$$\text{결합사후밀도함수: } \pi(\mu, \sigma^2 | x) \propto (\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)'(x - \mu) - \frac{1}{2}(\mu - \mu_0)'(\mu - \mu_0)\right) \sigma^{-a-1} e^{-\frac{b}{\sigma^2}}$$

$$\text{표본 추출: } \theta^{\text{prop}} \sim N(\theta^{\text{curr}}, \delta^2 I_2), \alpha = \frac{\pi(\theta^{\text{prop}} | x)}{\pi(\theta^{\text{curr}} | x)} \times \frac{q(\theta^{\text{curr}} | x)}{q(\theta^{\text{prop}} | x)}$$

$$\theta^{\text{next}} = \begin{cases} \theta^{\text{prop}} & \text{if } u \leq \alpha \\ \theta^{\text{curr}} & \text{if } u > \alpha \end{cases}, u \sim U(0, 1) \Rightarrow \theta^{\text{curr}} = \theta^{\text{next}}$$

예제 4.3 에 대한 MH 알고리즘에서 $\theta = (\mu, \log(\sigma^2))$ 의 후보표본을 다변량 정규분포 $N(\theta^{\text{curr}}, \delta^2 I_2)$ 에서 생성하여 θ 의 원소 모두에 같은 보폭(표준편차)을 사용하였다. 그런데 θ 의 원소마다 분산이 다를 수 있고 따라서 적절한 보폭이 다를 수 있으므로 $N(\theta^{\text{curr}}, \text{diag}(\delta_1^2, \delta_2^2))$ 을 고려하여 보자. $\text{diag}(\delta_1^2, \delta_2^2)$ 은 대각원소 δ_1^2, δ_2^2 를 갖는 대각행렬이다. 적절한 δ_1^2, δ_2^2 값을 찾기 위하여 다음의 과정을 실행한다.

- 1) 4.2 절의 알고리즘을 짧게 실행하여 $\theta = (\mu, \log(\sigma^2))$ 의 대략적인 추정치와 분산을 추정한다.
- 2) 추정된 분산에 2.4 를 곱하여 δ_1^2, δ_2^2 값으로 정한다.
- 3) 4.2 절의 알고리즘을 변형하여 $N(\theta^{\text{curr}}, \text{diag}(\delta_1^2, \delta_2^2))$ 로부터 후보표본을 추출하는 MH 알고리즘을 코딩한다.
- 4) MH 알고리즘을 짧게 수행한 후 후보표본의 채택확률을 계산하여 본다.
- 5) 채택확률이 약 24%가 되도록 δ_1^2, δ_2^2 를 조정한다.
- 6) 최종 선택된 δ_1^2, δ_2^2 를 사용하여 MH 를 충분히 길게 수행한 후 수렴속도, 효율, 사후추론 결과를 4.2 절의 공통분산을 사용하였을 경우와 비교하라.

RESULT

- 1) $\theta = (\mu, \log(\sigma^2))$ 에 대한 MH 알고리즘을 2000 번(burn-in: 500, Number of chain: 3) 수행한 후 대략적인 추정치와 분산은 다음과 같다. ($\delta = 1$)

	μ	$\log(\sigma^2)$
추정치	15.133	3.183
분산	2.32	0.227

- 2) 계산된 δ_1^2, δ_2^2

δ_1^2	δ_2^2
5.567	0.544

- 3) 기존의 알고리즘을 변형하여 후보 표본을 추출한다. 2000 번(burn-in: 500, Number of chain: 3)
- 4) 변형된 알고리즘을 사용하여 계산한 추정치와 분산은 다음과 같다.

	μ	$\log(\sigma^2)$
추정치	15.23	3.18
분산	2.467	0.223

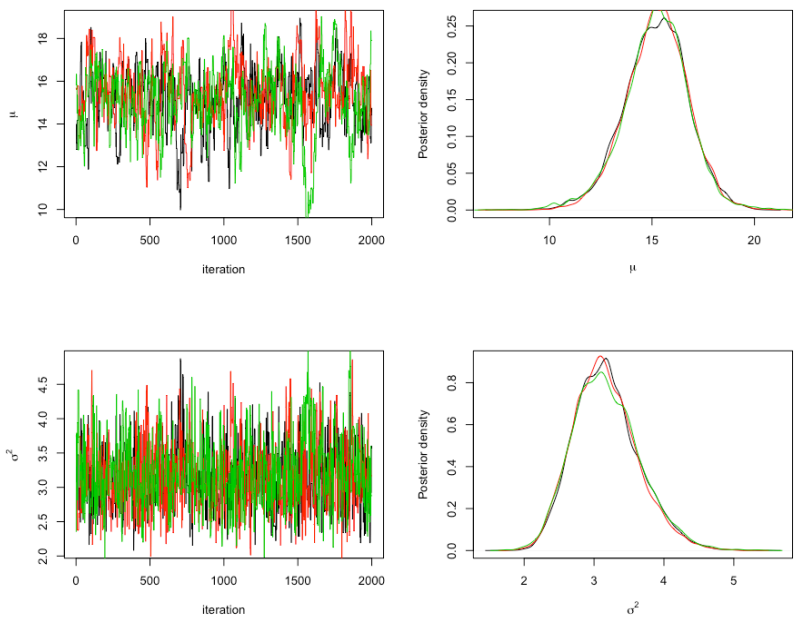
후보 표본의 채택 확률은 36.89%이다.

- 5) 채택 확률의 감소를 위하여 δ_1^2, δ_2^2 을 다음과 같이 설정하였다.

δ_1^2	δ_2^2
7.7	1.1

이때 후보 표본의 채택 확률은 27.19%이다.

- 6) 최종 선택된 δ_1^2, δ_2^2 를 사용하여 MH 알고리즘을 20000 번(burn-in: 5000) 수행하였다. 비교를 위하여 기존 알고리즘의 결과를 정리하면 다음과 같다.



위의 그래프는 추정 모수 별 처음 2000 개의 경로 그림 및 전체 표본에 대한 사후밀도함수이다.
Gelman 상수를 계산해 보았을 때, 모든 값들이 1.1 을 넘지 않기에 수렴하였음을 알 수 있다.

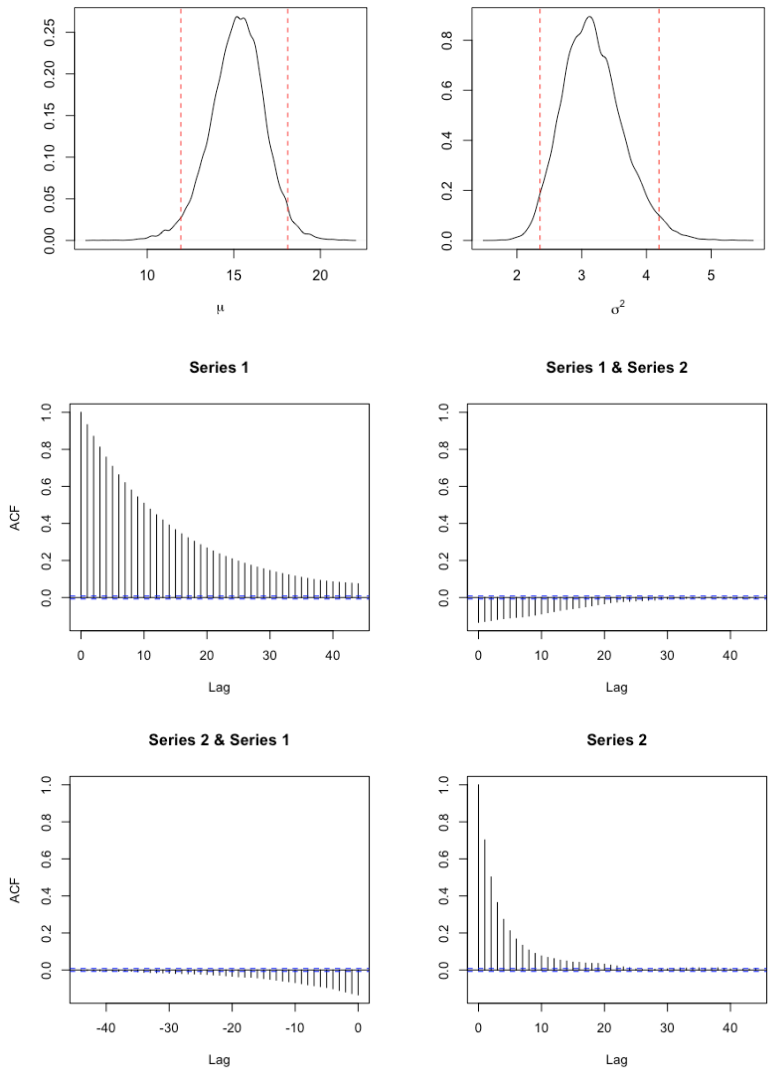
	Point est.	Upper C.I.
μ	1	1
$\log(\sigma^2)$	1	1

채택 확률은 40.08%이다.

추정치와 그 분산은 다음과 같다.

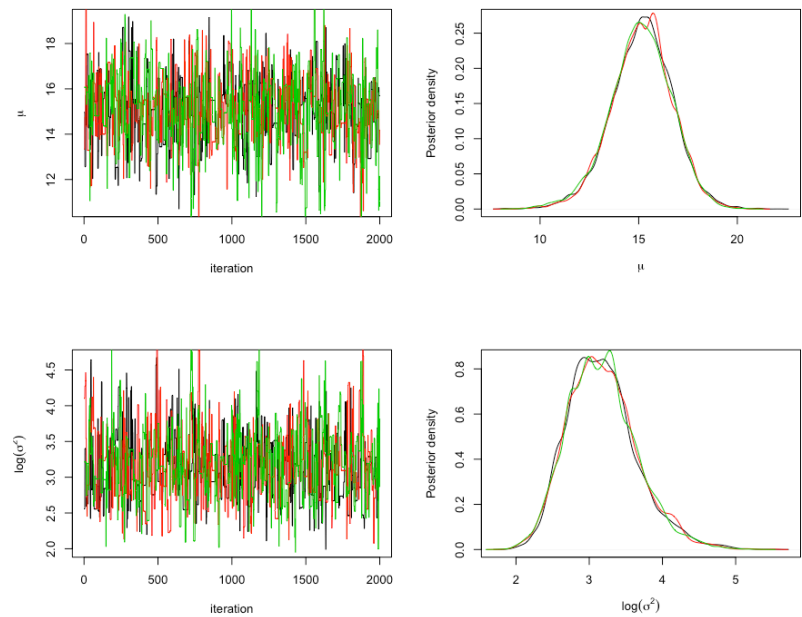
	μ	$\log(\sigma^2)$
추정치	15.206	3.173
분산	2.433	0.219

전체 샘플링 데이터에 대한 사후밀도함수와 HPD 값은 다음과 같다.



그래프를 살펴보았을 때, μ 에 대한 correlation 이 존재함을 알 수 있다.

채택된 δ_1^2, δ_2^2 를 사용하여 (이분산)MH 알고리즘을 수행한다.



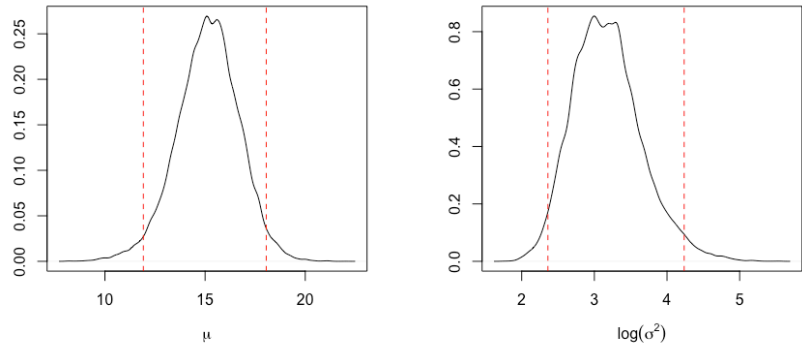
위의 그래프는 추정 모수 별 처음 2000 개의 경로 그림 및 전체 표본에 대한 사후밀도함수이다. Gelman 상수를 계산해 보았을 때, 모든 값들이 1.1 을 넘지 않기에 수렴하였음을 알 수 있다.

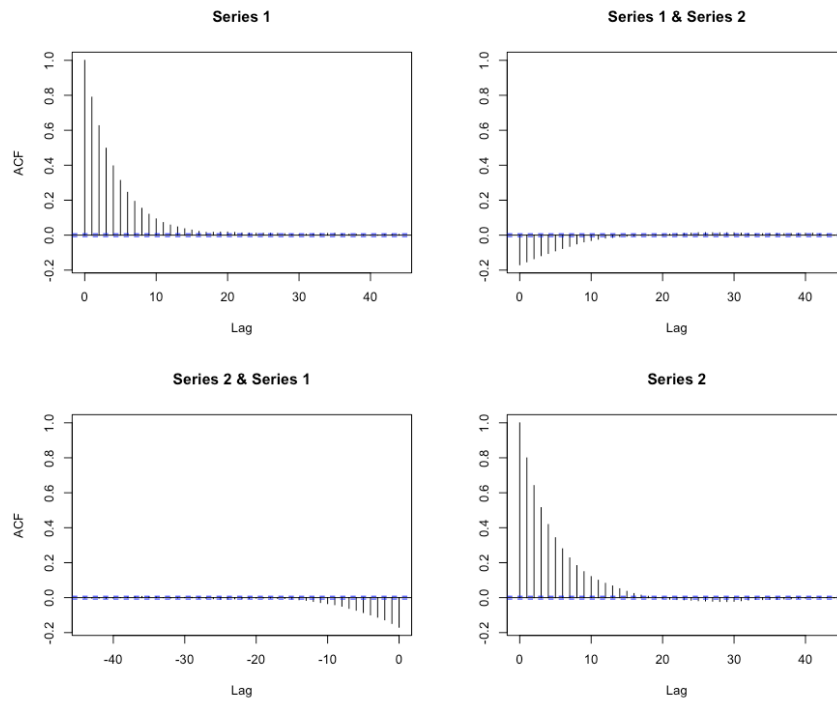
	Point est.	Upper C.I.
μ	1	1
$\log(\sigma^2)$	1	1

채택 확률은 24.68%이다.

추정치와 그 분산은 다음과 같다.

	μ	$\log(\sigma^2)$
추정치	15.143	3.178
분산	2.409	0.226





등분산 MH 알고리즘에 비해 이분산 MH 알고리즘을 사용하여 추출한 표본들이 correlation 이 적다는 것을 알 수 있다. 또한 추정치의 분산이 미세하게 더 작다. 그러나 사후밀도함수를 그려보았을 때, 이분산 MH 알고리즘 사용 시 그래프가 좀 더 불안정한것을 관찰할 수 있다. 두 번의 시뮬레이션 모두 많은 반복 수와 burn-in 을 사용하기에 수렴 속도를 확인하기에는 어려움이 있다. 그럼에도 이분산 MH 알고리즘의 표본추출이 좀 더 넓은 범위로 일어나는 양상을 보인다(좀 더 수렴속도가 빨라 보임). 두 알고리즘을 사용하여 추정된 $\mu, \log(\sigma^2)$ 의 값은 비슷하다.