STA 602 Lab 8

Ryan Tang

08 November, 2022

Exercise 1

$$\begin{aligned}
y_{i1} &= x_{i1} - \theta_1 && (1)\\
y_{i2} &= x_{i2} - \theta_2 && (2)\\
\lambda &= \frac{1}{1-\rho^2} && (3)\\
p(\theta_1|X,\rho,\theta_2) &\propto \exp[-\frac{\lambda}{2}[\lambda^{-1}\theta_1^2 + \sum_i^N (y_{i1}^2 + y_{i2}^2 - 2\rho y_{i1}y_{i2})]] && (4)\\
&\propto \exp[-\frac{\lambda}{2}[\lambda^{-1}\theta_1^2 + n\theta_1^2 - 2\theta_1 n\bar{X}_1 + 2\rho\theta_1 n\bar{X}_2 - 2\rho n\theta_1\theta_2]] && (5)\\
&\propto \exp[-\frac{\lambda}{2}[\theta_1^2(n+\lambda^{-1}) - 2n\theta_1(\bar{X}_1 - \rho(\bar{X}_2 - \theta_2))]] && (6)\\
&\propto \exp[-\frac{n\lambda+1}{2}[\theta_1^2 - 2\theta_1\frac{n\lambda}{n\lambda+1}(\bar{X}_1 - \rho(\bar{X}_2 - \theta_2)]] && (7)\\
\sigma_{1n}^2 &= (n\lambda+1)^{-1} = \lambda_{1n}^{-1} && (8)\\
\mu_{1n} &= \frac{n\lambda}{\lambda_{1n}}(\bar{X}_1 - \rho(\bar{X}_2 - \theta_2)) && (9)\\
\theta_1|X,\rho,\theta_2 &\sim \mathcal{N}(\theta_1|\mu_{1n},\sigma_{1n}^2) && (10)
\end{aligned}$$

Exercise 2

$$\begin{aligned}
\sigma_{2n}^2 &= (n\lambda+1)^{-1} = \lambda_{2n}^{-1} = \lambda_{1n}^{-1} && (11)\\
\mu_{2n} &= \frac{n\lambda}{\lambda_{2n}}(\bar{X}_2 - \rho(\bar{X}_1 - \theta_1)) && (12)\\
\theta_2|X,\rho,\theta_1 &\sim \mathcal{N}(\theta_2|\mu_{2n},\sigma_{2n}^2) && (13)
\end{aligned}$$

Exercise 3

We can see from the full conditionals, the posterior mean θ_1 and θ_2 are a linear function of observed mean, and scalled back by the ρ as a coefficient.

Exercise 4

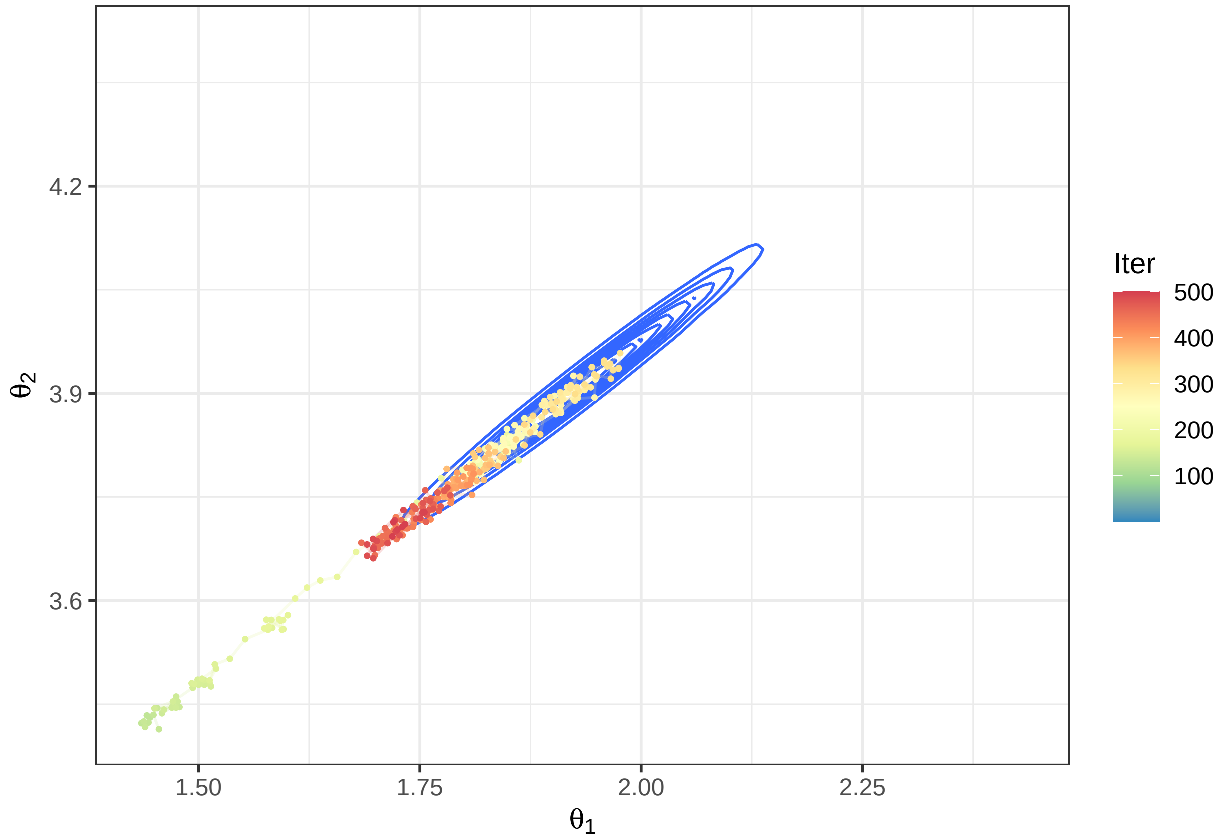
The Gibbs implementation

```
normal_gibbs_sampler <- function(S, X, rho) {
  N = dim(X)[1]
```

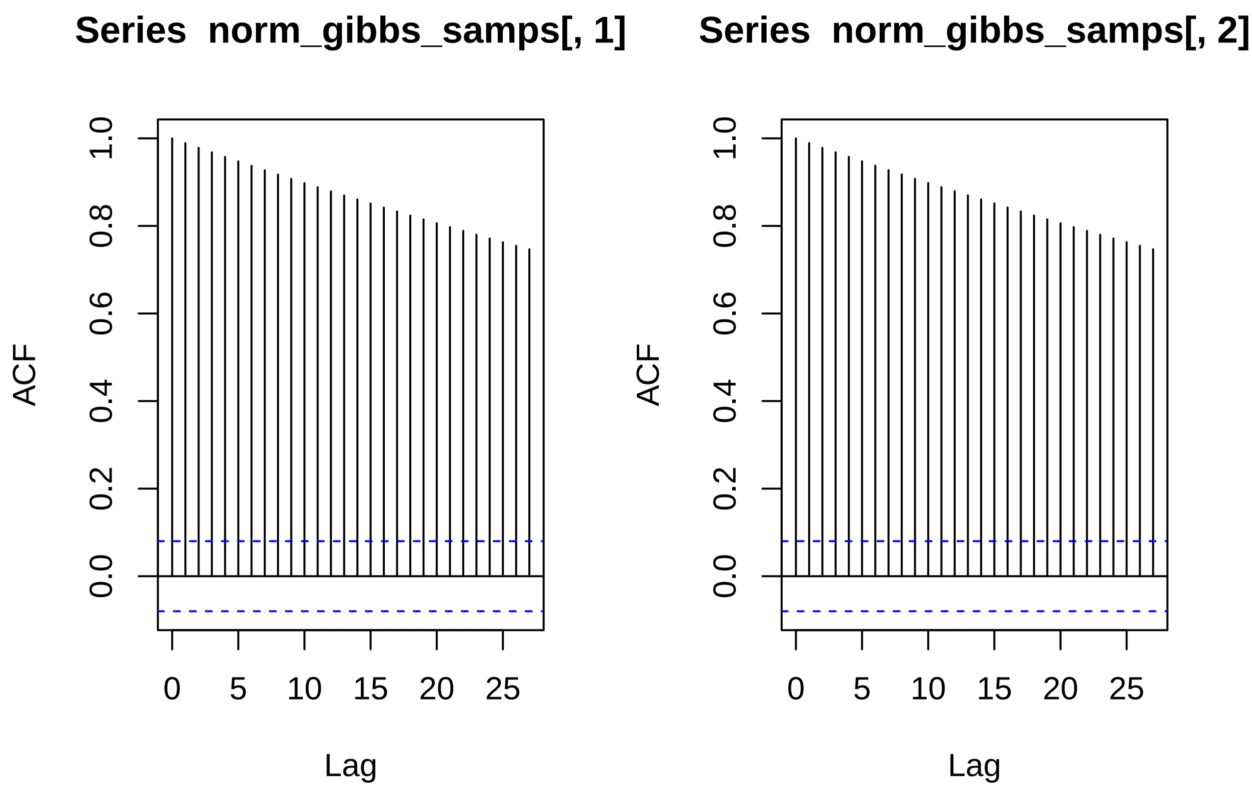
```
  X_mu1 = mean(X[, 1])
  X_mu2 = mean(X[, 2])

  theta1 = 0
  theta2 = 0
  samples = matrix(nrow=S, ncol=2)
  for (s in 1:S) {
    lambda = 1 / (1 - rho^2)
    theta1 = rnorm(1, mean=(N*lambda)/(N*lambda+1) * (X_mu1 - rho*(X_mu2 - theta2)), sd=sqrt(1/(N*lambda
    theta2 = rnorm(1, mean=(N*lambda)/(N*lambda+1) * (X_mu2 - rho*(X_mu1 - theta1)), sd=sqrt(1/(N*lambda
    samples[s, ] = c(theta1, theta2)
  }

  return(samples)
}
```

Using the Gibbs sampler.

```
n <- 100
rho <- 0.995
X <- MASS::mvrnorm(n = n, mu = c(2, 4), Sigma = matrix(c(1, rho, rho, 1), nrow = 2))
Sigma_post <- matrix(((1-rho^2)/((n+1-rho^2)^2 - (n^2)*(rho^2)))*c(n+1-rho^2, n*rho, n*rho, n+1-rho^2),
mu_post <- n*Sigma_post%*%matrix(c(1/(1-rho^2), -rho/(1-rho^2),
                                   -rho/(1-rho^2), 1/(1-rho^2)),
                                   nrow = 2)%*%colMeans(X)
norm_gibbs_samps <- normal_gibbs_sampler(600, X, rho)
#
true_post <- MASS::mvrnorm(n = 100000,
                           mu = mu_post,
                           Sigma = Sigma_post)
data.frame(norm_gibbs_samps) %>%
  magrittr::set_colnames(c("theta_1", "theta_2")) %>%
  dplyr::mutate(iter = 1:n()) %>%
  dplyr::filter(iter > 100) %>%
  dplyr::mutate(iter = 1:n()) %>%
  ggplot2::ggplot() +
  geom_density2d(data = data.frame(true_post) %>%
                        magrittr::set_colnames(c("true_1", "true_2")),
                 aes(x = true_1, y = true_2)) +
  geom_path(aes(x = theta_1, y = theta_2, colour = iter), alpha = 0.2, size = 0.5) +
  geom_point(aes(x = theta_1, y = theta_2, colour = iter), size = 0.5) +
  scale_color_distiller(palette = "Spectral", name = "Iter") +
  labs(x = expression(theta[1]), y = expression(theta[2])) +
  xlim(c(mu_post[1] - 0.5, mu_post[1] + 0.5)) +
  ylim(c(mu_post[2] - 0.5, mu_post[2] + 0.5))
```

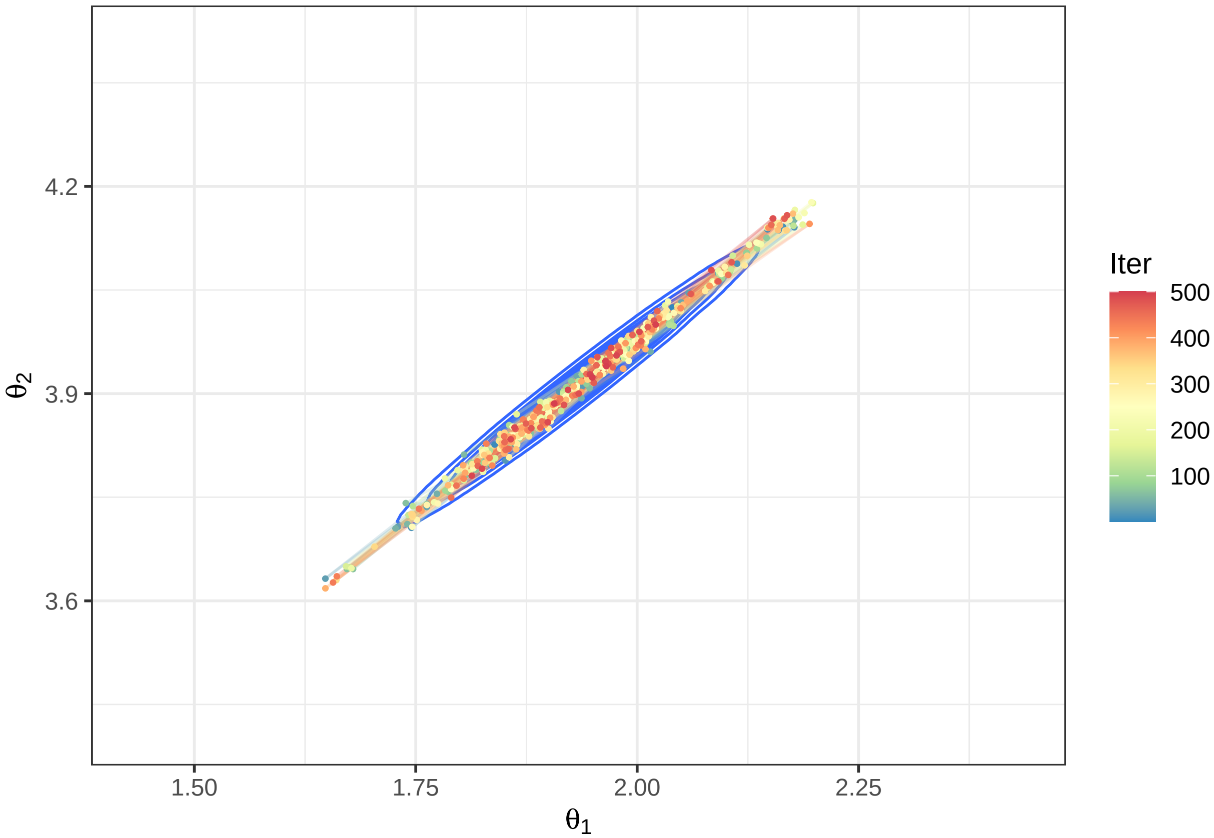



```r
par(mfrow = c(1,2))
acf(norm_gibbs_samps[,1])
acf(norm_gibbs_samps[,2])
```

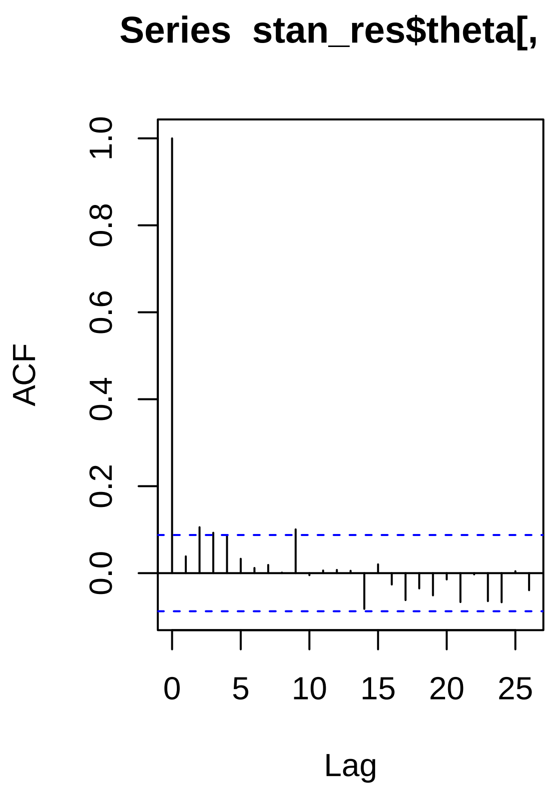


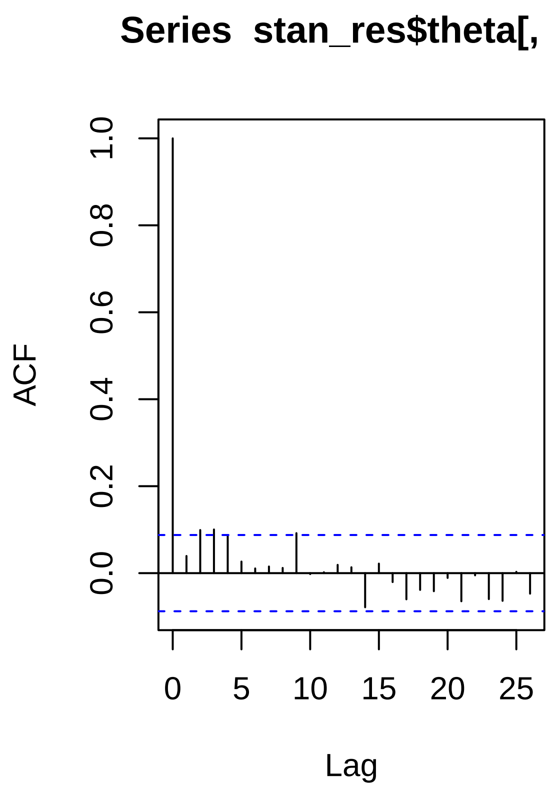
Using the HMC sampler from STAN.

```
# Generates samples by HMC
stan_res <- rstan::stan("lab-08-hmc_norm_example.stan", data = list(X = X,
                                                              N = nrow(X),
                                                              Sigma = matrix(c(1, rho, rho, 1), nrow = 2
                        chains = 1, iter = 600, warmup = 100, verbose = F, refresh = 0) %>%
            rstan::extract()
#
data.frame(stan_res$theta) %>%
  magrittr::set_colnames(c("theta_1", "theta_2")) %>%
  dplyr::mutate(iter = 1:n()) %>%
  ggplot2::ggplot() +
  geom_density2d(data = data.frame(true_post) %>%
                        magrittr::set_colnames(c("true_1", "true_2")),
                 aes(x = true_1, y = true_2)) +
  geom_path(aes(x = theta_1, y = theta_2, colour = iter), alpha = 0.2, size = 0.5) +
  geom_point(aes(x = theta_1, y = theta_2, colour = iter), size = 0.5) +
  scale_color_distiller(palette = "Spectral", name = "Iter") +
  labs(x = expression(theta[1]), y = expression(theta[2])) +
  xlim(c(mu_post[1] - 0.5, mu_post[1] + 0.5)) +
  ylim(c(mu_post[2] - 0.5, mu_post[2] + 0.5))
```




```r
par(mfrow = c(1,2))
acf(stan_res$theta[,1])
acf(stan_res$theta[,2])
```





Exercise 5

Given $\rho = 0.995$, Gibbs sampler had a super sticky chain. Many samples are outside of the high probability area. And the top-right corner was not well explored either. ACF plot also shows there are strong auto-correlation between samples.

Exercise 6

The posterior mean μ_{1n} for θ_1 is given $\propto \bar{X}_1 - \rho(\bar{X}_2 - \theta_2)$. When ρ is large, the posterior mean barely moves and is artificially stuck at somewhere close to 0. And the same effect applies to θ_2 because θ_1 is not moving; hence the stickiness spirals.