HW1

Seungjun Lee

2022 9 29

1-(a)

As we know, linear transformation of multivariate normal distribution is also multivariate normal distribution. The mean and variance are as follow.

$$E(Y) = \mu + L * E[Z] = \mu$$

$$Var(Y) = LVar(Z)L^T = LL^T = \Sigma$$

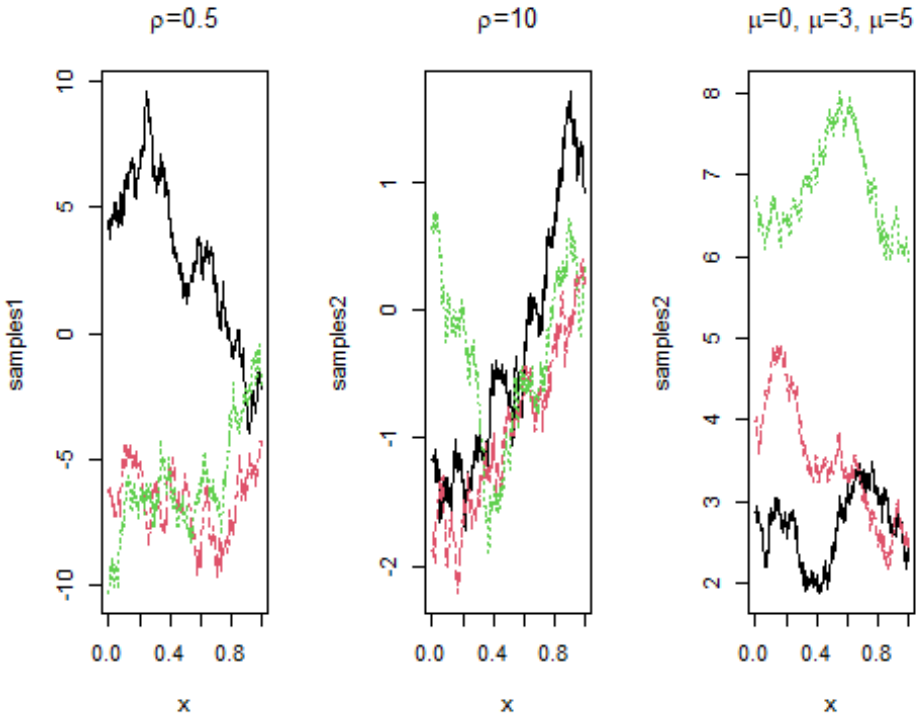
1-(b)

```
# 1-(b)

sim_y = function(mu,sigma){
  L =t(chol(sigma)) # lower triangle matrix
  n = length(mu)
  Z = rnorm(n,0,1)
  Y = mu + L%*%Z
  return(Y)
}
```

1-(c)

We consider an exponential covariance function.

$$C(h) = \sigma^2 \exp(-\| h \| / \rho)$$



2-(a)

```
load("C:/Users/user/Desktop/jun/SpatioTemp/hw1/CAtemps.RData")


head(CAtemp)
```

```
##                   coordinates  avgtemp  elevation
## 040136 (-116.7667, 32.83333) 63.23431  564.32227
## 040161   (-120.55, 41.48333) 46.18375 1346.82019
## 040212 (-122.4333, 38.56667) 56.84736  566.09875
## 040232 (-121.7333, 37.98333) 60.08181   26.57045
## 040343 (-118.8167, 36.48333) 62.97833 1005.60638
## 040379    (-119.5, 37.08333) 60.46861  679.55133
```

```
data = data.frame(cbind(coordinates(CAtemp),CAtemp$avgtemp,CAtemp$elevation))
sum(is.na(data)) # no missing data
```

```
## [1] 0
```

```
colnames(data) = c("lon","lat","avgtemp","elevation")
fit = lm(avgtemp ~ lon + lat + elevation, data=data)
# summary(fit)
coeff = fit$coefficients  # answer
res = data$avgtemp - fit$fitted.values
coeff
```

```
##   (Intercept)           lon           lat     elevation
## 321.511433492   2.324104683   0.564680460  -0.009647649
```

```
# plot.point.ref(CAtemp, res)
```

```
ploteqc <- function(spobj, z, breaks, ...){
  pal <- tim.colors(length(breaks)-1)
  fb <- classIntervals(z, n = length(pal),
                       style = "fixed", fixedBreaks = breaks)
  col <- findColours(fb, pal)
  plot(spobj, col = col, ...)
  image.plot(legend.only = TRUE, zlim = range(breaks), col = pal)
}
```

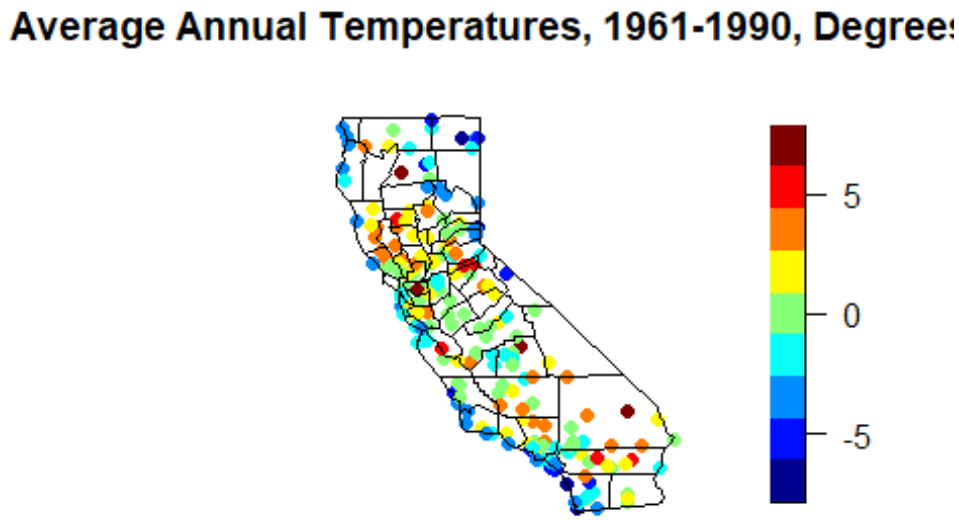
```
range(res)
```

```
## [1] -6.304300  6.953582
```

```
breaks = seq(-7, 7, length.out = 10)
ploteqc(CAtemp, res, breaks, pch = 19) # answer
```

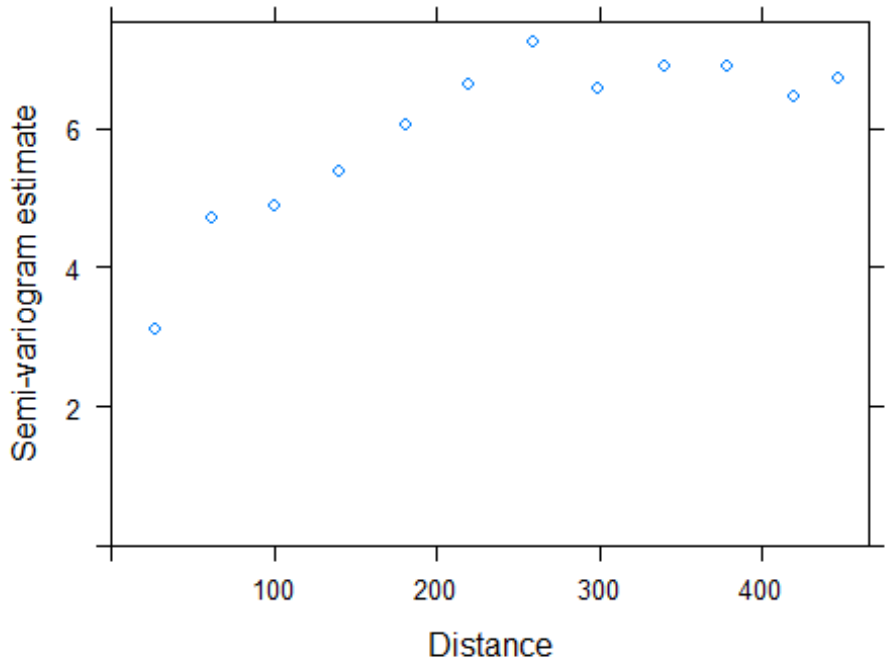
```
## Warning in wkt(obj): CRS object has no comment
```

```
map("county", region = "california", add = TRUE)
title(main = "Average Annual Temperatures, 1961-1990, Degrees F")
```

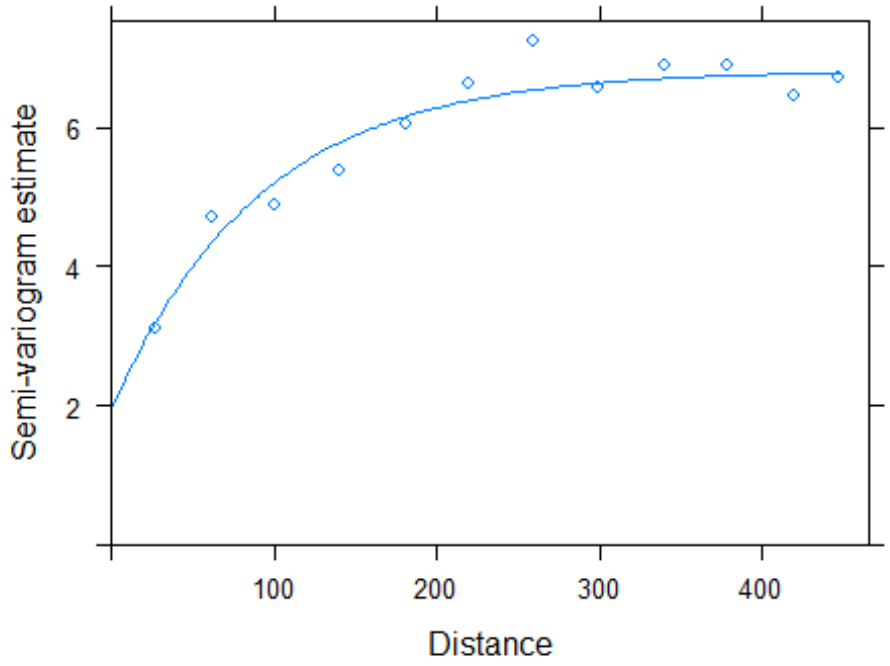


2-(b)

```
CAtemp$res = res

vg = variogram(res ~ 1, data = CAtemp, width=40)

## Warning in wkt(obj): CRS object has no comment

plot(vg, xlab = "Distance", ylab = "Semi-variogram estimate", width=5)
```




```
# vgangle = variogram(res ~ 1, data = CAtemp, alpha = c(0, 45, 90, 135))
# plot(vgangle, xlab = "Distance", ylab = "Semi-variogram estimate")
# # I think there are no big differences..

vgmodel = vgm(1, "Exp", 200, 0.05)
fitvg = fit.variogram(vg, vgmodel)
plot(vg, fitvg, xlab = "Distance", ylab = "Semi-variogram estimate")
```




```
print(fitvg)

##   model    psill    range
## 1   Nug 1.932258  0.00000
## 2   Exp 4.881704 90.56163

s2.hat = fitvg$psill[2]
rho.hat = fitvg$range[2]
tau2.hat = fitvg$psill[1]
```

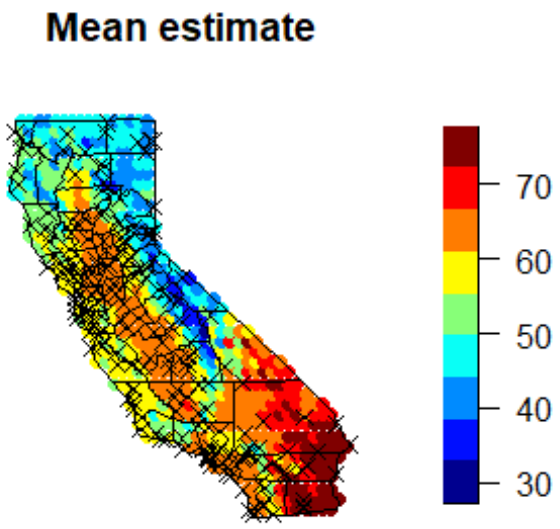
2-(c)

```
# 2-(c)
head(data)

##         lon      lat  avgtemp  elevation
## 1 -116.7667 32.83333 63.23431  564.32227
## 2 -120.5500 41.48333 46.18375 1346.82019
## 3 -122.4333 38.56667 56.84736  566.09875
## 4 -121.7333 37.98333 60.08181   26.57045
## 5 -118.8167 36.48333 62.97833 1005.60638
## 6 -119.5000 37.08333 60.46861  679.55133

data2 = cbind(coordinates(CAtemp),CAtemp$elevation)
colnames(data2) = c("lon","lat","elevation")
X = as.matrix(cbind(1,data2))
```

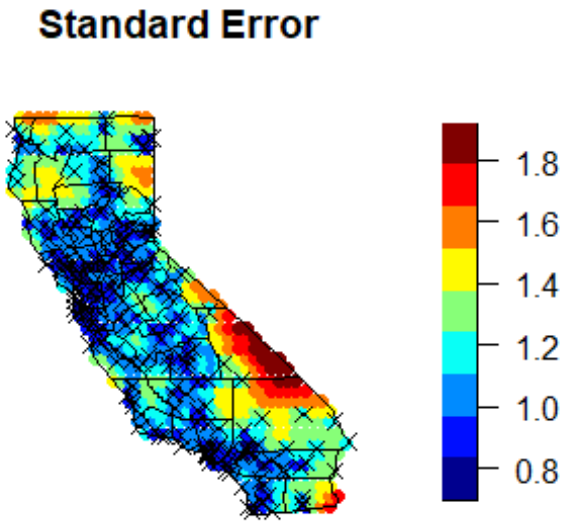
```
dist = rdist.earth(coordinates(CAtemp))
covmat = exp(-dist/rho.hat)*s2.hat
diag(covmat) = s2.hat
nugget = diag(nrow(dist))*tau2.hat
Sigma = covmat + nugget
Sigma_inv = solve(covmat + nugget)
y = CAtemp$avgtemp
beta_gls = solve(t(X)%*% Sigma_inv %*% X) %*% t(X) %*% Sigma_inv %*% y

rownames(beta_gls)[1] = "intercept"
beta_gls # answer

##                     [,1]
## intercept 355.096049673
## lon         2.619671554
## lat         0.571798967
## elevation  -0.009045026
```

2-(d)

```
#2-(d)
dd = rdist.earth(coordinates(CAtemp),coordinates(CAgrid))
gamma = s2.hat * exp(-dd/rho.hat)
Xs = cbind(1,coordinates(CAgrid),CAgrid$elevation)

ypred = Xs %*% beta_gls + t(gamma) %*% solve(Sigma) %*%(y - X %*% beta_gls)

range(ypred)

## [1] 33.26956 73.72527

breaks = seq(30, 75, length.out = 10)
ploteqc(CAgrid, ypred, breaks, pch = 19) # answer

## Warning in wkt(obj): CRS object has no comment

map("county", region = "california", add = TRUE)
points(coordinates(CAtemp),pch=4,cex=1)
title(main = "Mean estimate")
```




```
b = t(Xs) - t(X) %*% solve(Sigma) %*% gamma
vpred = s2.hat - diag(t(gamma) %*% Sigma_inv%*%gamma + t(b) %*% solve(t(X) %*%
 Sigma_inv %*% X ) %*% b)


sepred = sqrt(vpred)

range(sepred)

## [1] 0.7643745 1.8528903

breaks = seq(range(sepred)[1],range(sepred)[2], length.out = 10)
ploteqc(CAgrid, sepred, breaks, pch = 19) # answer

## Warning in wkt(obj): CRS object has no comment

map("county", region = "california", add = TRUE)
points(coordinates(CAtemp),pch=4,cex=1)
title(main = "Standard Error")
```


Standard Error
1.8
1.6
1.4
1.2
1.0
0.8