机器学习作业

邵智轩　　1400012141　　物理学院

零均值高斯 ARMA(1,1) 模型：

$$X_t = aX_{t-1} + e_t + be_{t-1}, \quad t = 1, 2, ..., T \tag{1}$$

其中 $0 < |a| < 1$，$0 < |b| < 1$，$-a \neq b$，$\{e_t\}$ 为 i.i.d. 的 $N(0, \sigma^2)$ 分布。

给定观测 $x_1, x_2, ..., x_T$，写出在 $x_0 = 0$ 且 $e_0 = 0$ 条件下的条件对数似然函数

$$l(a, b, \sigma^2) = -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T} e_t^2 \tag{2}$$

（忽略常数项 $-\frac{T}{2}\log(2\pi)$。其中，由可逆性可迭代解得 e_t（含参数 a、b）：

$$e_t = X_t - aX_{t-1} - be_{t-1}, \quad t = 1, 2, ..., T \tag{3}$$

编程部分

```
negLog <- function(theta, x) {
    # theta[1]=a, theta[2]=b, theta[3]=log(sigma^2)
    N <- length(x)
    a<-theta[1]
    b<-theta[2]
    s2<-exp(theta[3])
    e_lag <- 0
    x_lag <- 0
    RSS <- 0
    e <- numeric()
```

```
    for (x_i in x) {
        e <- x_i - a * x_lag - b * e_lag
        RSS <- RSS + e ^ 2
        e_lag <- e
        x_lag <- x_i
    }
    RSS / (2 * s2) + N / 2 * log(s2)
}

ARMA11 <- function(seed=1, theta, N) {
    set.seed(seed)
    e <- rnorm(N, sd = sqrt(theta[3]))
    x <- e
    for (i in 2:N) {
        x[i] <- theta[1] * x[i - 1] + e[i] + theta[2] * e[i - 1]
    }
    x
}

theta.real<-c(0.5,0.7,1)
theta.MLE<-t(sapply(1:1000, function(seed){
  ARMAData<-ARMA11(seed=seed, theta=theta.real,N=100)
  ores<-optim(c(0,0,0),negLog,x=ARMAData)
  theta.MLE<-ores$par
}))
theta.MLE[,3]<-exp(theta.MLE[,3])
summary(theta.MLE)
```

```
##        V1                V2               V3
##  Min.   :0.08837   Min.   :0.4171   Min.   :0.6443
##  1st Qu.:0.42649   1st Qu.:0.6566   1st Qu.:0.8763
##  Median :0.49509   Median :0.7129   Median :0.9762
##  Mean   :0.49021   Mean   :0.7101   Mean   :0.9817
```

```
##  3rd Qu.:0.56190   3rd Qu.:0.7666   3rd Qu.:1.0691
##  Max.   :0.74204   Max.   :1.0103   Max.   :1.4413
```

```
cat(' 真实 a=', theta.real[1], 'MLE 估计 =', mean(theta.MLE[,1]),
    ' 估计的标准误 =', sd(theta.MLE[,1]), '\n')
```

```
## 真实a= 0.5 MLE估计= 0.490213 估计的标准误= 0.1031492
```

```
cat(' 真实 b=', theta.real[2], 'MLE 估计 =', mean(theta.MLE[,2]),
    ' 估计的标准误 =', sd(theta.MLE[,2]), '\n')
```

```
## 真实b= 0.7 MLE估计= 0.7101126 估计的标准误= 0.08505276
```

```
cat(' 真实 s2=', theta.real[3], 'MLE 估计 =', mean(theta.MLE[,3]),
    ' 估计的标准误 =', sd(theta.MLE[,3]), '\n')
```

```
## 真实s2= 1 MLE估计= 0.9817078 估计的标准误= 0.1414473
```