

Multivariate Normal Distribution Exploration

Joel S Steele, PhD

August 4, 2016

```
library(knitr)
opts_chunk$set(comment='')
```

```
library(mvtnorm)
sigma <- matrix(c(4,2,2,3), ncol=2)
x <- rmvnorm(n=500, mean=c(1,2), sigma=sigma)
colMeans(x)
```

```
[1] 1.098951 2.064273
```

```
var(x)
```

```
      [,1]      [,2]
[1,] 3.916565 2.112068
[2,] 2.112068 3.165835
```

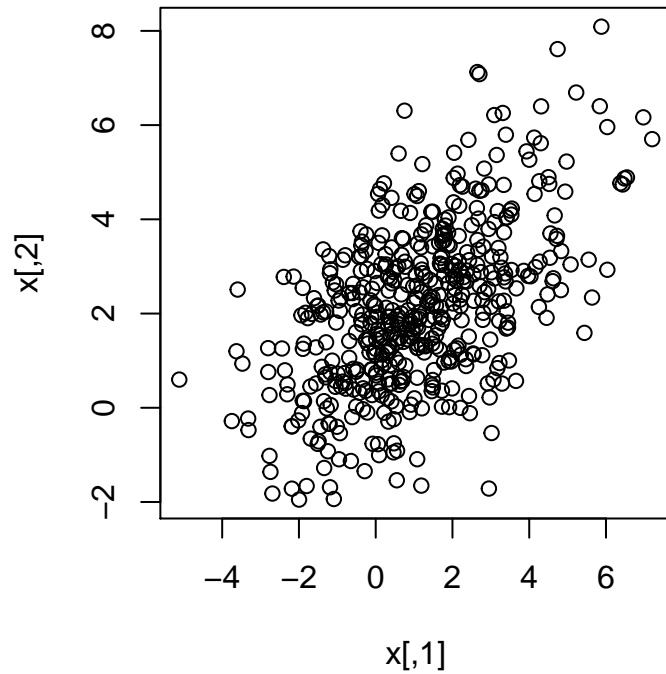
```
x <- rmvnorm(n=500, mean=c(1,2), sigma=sigma, method="chol")
colMeans(x)
```

```
[1] 1.088509 2.121251
```

```
var(x)
```

```
      [,1]      [,2]
[1,] 3.929686 1.882966
[2,] 1.882966 2.986817
```

```
op = par(pty='s')
plot(x)
```



```
par(op)
```

Density estimation

```
mvnllik = function(dat,par){
  mu1 = par[1]
  mu2 = par[2]
  sig1 = par[3]
  sig2 = par[4]
  cv12 = par[5]
  ll = -sum(
    dmvnorm(x=dat,
            mean=c(mu1,mu2),
            sigma=matrix(c(sig1,cv12,cv12,sig2),nrow=2),
            log=FALSE)
  )
  return(ll)
}

#res = optim(c(1,2,4,2,2.75),fn=mvnllik,dat=x,control=list(trace=TRUE, REPORT=1))
res = optim(c(1,2,4,2,2.75),fn=mvnllik,dat=x,method='CG')
res$par
```

```
[1] 1.024984 1.958659 3.969831 1.973642 2.796086
```

```
res2 = nlm(b,c(1,2,4,2,2.75),mvnllik, dat=x)
res2$par
```

```
[1] 1.059446 1.901769 3.952258 1.980627 2.797847
```

```
kable(cbind('SampEst'=c(colMeans(x),var(x)[c(1,4,2)]), 'optim'=res$par, 'nlminb'=res2$par))
```

SampEst	optim	nlminb
1.088509	1.024984	1.059446
2.121251	1.958659	1.901769
3.929686	3.969831	3.952258
2.986817	1.973642	1.980627
1.882966	2.796086	2.797847

```
# attempt 1 at a direct optimization method for MVN distribution
# let's start out with a bivariate distribution
# we will also be optimizing the log-likelihood
ll_mvn = function(parm,dat){
  n = nrow(dat)
  S = var(dat)
  # parameters for our bivariate normal distribution
  mu1 = parm[1]
  mu2 = parm[2]
  muvec = c(mu1,mu2)
  sigma = matrix(c(parm[3],parm[5],parm[5],parm[4]),nrow=2)
  # cholesky inverse of sigma
  # u = chol(sigma)
  # uinv = solve(u)
  # sinv = uinv *t(uinv)
  sinv = solve(sigma)
  # log likelihood
  ll = n*(log(det(sigma)) + sum(diag(sinv*S)))
  return(ll)
}

#res3 = optim(c(1,2,4,3,2),fn=ll_mvn,control=list("fnscale"=-1),dat=x)
```