

Stat 471: Lecture 5

Multivariate normal distributions

Moo K. Chung
mchung@stat.wisc.edu

September 15, 2003

- For given $z_1, z_2 \sim^{i.i.d.} N(0, 1)$, we define a random vector $z = (z_1, z_2)'$. Let $H = (h_{ij})$. Then we define a multivariate normal distribution as $w = Hz \sim N(0, HH')$, where HH' is called the *covariance matrix*.

- The covariance matrix HH' is computed in the following way.

$$\text{Var}(Hz) = E[(Hz)(Hz)'] = HH'$$

assuming the mean is 0. The covariance matrix is *symmetric positive definite*.

- Any symmetric positive definite matrix $V = (v_{ij})$ can be factored as $V = HH'$, where $H = (h_{ij})$ is the lower triangular matrix and H' is the upper triangular matrix.

```
>>V=[2 1
      1 2]
>>Htrans=chol(V)
Htrans =
      1.4142      0.7071
           0      1.2247
>> Htrans'*Htrans
ans =
      2.0000      1.0000
      1.0000      2.0000
```

It can be shown that $h_{11} = \sqrt{v_{11}}$ and $h_{i1} = v_{i1}/h_{11}$ for $i = 2, \dots, n$ (HW 2).

- Generate 1000 multivariate normal random vectors with zero mean and covariance $V = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. *Solution.* From HW 1, I assume that everyone will have his/her own version of standard normal random number generator called `snrnd`. I assume that `snrnd` has one argument n , the number of random numbers it is generating.

```
>>z1 = snrnd(1000); z2=snrnd(1000);
>>w=Htrans'*[z1 z2]';
>> w(:,1:3)
ans =
      0.1772      0.4068     -1.6214
     -2.4402     -0.7073     -0.5949
```

- For given bivariate data w_i , test if the data follows multivariate normal. You may use χ^2 goodness-of-fit test but a simpler approach would be to check if given bivariate data can be generated via linear transform $w = Hz + \mu, z \sim N(0, I)$. First estimate the sample covariance matrix:

$$\hat{V} = \frac{1}{n} \sum_{i=1}^n (w_i - \bar{w})(w_i - \bar{w})',$$

where $\bar{w}$ is the sample mean vector. The sample variance matrix is an estimator of unknown population covariance V . It can be computed using the *Kroneker tensor product* $\otimes$. For $A = (a_{ij})$ and B , $A \otimes B = (a_{ij}B)$.

```
>>barw=mean(w,2); temp=w-kron(ones(1,1000),barw);
>>hatV=temp*temp'/1000;
hatV =
      2.0982      1.0232
      1.0232      1.9711
```

Let $\hat{V} = \hat{H}\hat{H}'$ be the Cholesky factors. Then $z = \hat{H}^{-1}w$.

```
>> hatH=chol(hatV)'
hatH =
      1.4485          0
      0.7064      1.2133
>> inv(hatH)
ans =
      0.6904          0
     -0.4019      0.8242
>> newz=inv(hatH)*w;
>>qqplot(newz(1,:));qqplot(newz(2,:));
```