Covariance and correlation

```
>> load hospital.mat
>> X = [hospital.Weight hospital.BloodPressure]

X =

   176   124    93
   163   109    77
   131   125    83
   186   119    74
   172   136    93
   177   114    86
   ……..

>> C=cov(X)

C =

  706.0404   27.7879   41.0202
   27.7879   45.0622   23.8194
   41.0202   23.8194   48.0590

>> Y = [hospital.Age hospital.BloodPressure]

Y =

    38   124    93
    43   109    77
    49   119    74
    45   136    93
    48   114    86
……

>> D=corrcoef(Y)

D =

    1.0000    0.1341    0.0806
    0.1341    1.0000    0.5118
    0.0806    0.5118    1.0000
```

See also:

hospital.m file provided with this model answer

Corrcov() - convert covariance matrix to correlation matrix

R = corrcov(C) returns the correlation matrix R corresponding to the covariance matrix C.

[R,sigma] = corrcov(C) also returns sigma, a vector of standard deviations.

Compare Correlation Matrices Obtained by Two Different Methods

Compare the correlation matrix obtained by applying **corrcov** on a covariance matrix with the correlation matrix obtained by direct computation using **corrcoef** on an input matrix.

Load the hospital data set and create a matrix containing the **Weight** and **BloodPressure** measurements. Note that **hospital.BloodPressure has two columns of data**.

```
>> load hospital
>> X = [hospital.Weight hospital.BloodPressure];
>> C = cov(X)

C =

  706.0404   27.7879   41.0202
   27.7879   45.0622   23.8194
   41.0202   23.8194   48.0590

>> R1 = corrcov(C)

R1 =

    1.0000    0.1558    0.2227
    0.1558    1.0000    0.5118
    0.2227    0.5118    1.0000

>> R2 = corrcoef(X)

R2 =

    1.0000    0.1558    0.2227
    0.1558    1.0000    0.5118
    0.2227    0.5118    1.0000
```

Optional (advanced): Find Standard Deviations from Covariance Matrix

Find the vector **of standard deviations from the covariance matrix**, and show the relationship between the standard deviations and the covariance matrix.

Load the hospital data set and create a matrix containing the **Weight, BloodPressure,** and **Age** measurements. **Note that hospital.BloodPressure has two columns of data**.

```
>> load hospital
>> X = [hospital.Weight hospital.BloodPressure hospital.Age];
>> C = cov(X)
```

C =

```
  706.0404   27.7879   41.0202   17.5152
   27.7879   45.0622   23.8194    6.4966
   41.0202   23.8194   48.0590    4.0315
   17.5152    6.4966    4.0315   52.0622
```

C is square, symmetric, and positive semidefinite. The diagonal elements of C are the variances of the four variables in X.

Compute the correlation matrix and standard deviations of X from the covariance matrix C.

```
>> [R,s1] = corrcov(C)
```

R =

```
    1.0000    0.1558    0.2227    0.0914
    0.1558    1.0000    0.5118    0.1341
    0.2227    0.5118    1.0000    0.0806
    0.0914    0.1341    0.0806    1.0000
```

s1 =

```
   26.5714
    6.7128
    6.9325
    7.2154
```

Compute the square root of the diagonal elements in C, and then compare s1 with s2.

```
>> s2 = sqrt(diag(C))
```

s2 =

```
   26.5714
    6.7128
    6.9325
    7.2154
```

s1 and s2 are equal and correspond to the standard deviation of the variables in X.

Decision Trees

Please go to the example on data set **fisheriris** provided at:

https://uk.mathworks.com/help/stats/compactclassificationtree.predict.html?searchHighlight=predict&s_tid=doc_srchtitle#bst08bg-4