Generating random vectors with a given covariance matrix

BY

James Amundson

January 6, 2004

Ideal case

Given a ideal set of uncorrelated vectors $\{\rho^i\}$ with components evenly distributed between 0 and 1, we seek a transformation

$$r = A\rho$$

such that the set of vectors $\{r^i\}$ has covariance matrix

$$\langle r_i r_j \rangle = C_{\mathrm{ij}}.$$

We have

$$C = A\langle \rho_i \rho_j \rangle A^T,$$

or simply

$$C = \mathrm{AA}^T.$$

Since the covariance matrix is positive definite, the matrix A can be shown to be lower diagonal. Solving the equation above for A is known as the *Cholesky decomposition.* We can use the fact that C is symmetric to easily solve for the components of A.

```
Maxima 5.9.0.1cvs http://maxima.sourceforge.net
Using Lisp CMU Common Lisp 18d
Distributed under the GNU Public License. See the file COPYING.
Dedicated to the memory of William Schelter.
This is a development version of Maxima. The function bug_report()
provides bug reporting information.

(C1) load("f90.lisp")$
(C2) dim:6$
(C3) A:genmatrix(aa,dim,dim)$
(C4) for i:1 thru dim do for j:i+1 thru dim do A[i][j]:0$
(C5) AAt:A.transpose(A)$
(C6) C:genmatrix(cc,dim,dim)$
(C7) solns:[]$
(C8) for i:1 thru dim do for j:1 thru i do block( eq:AAt[i][j] = C[i][j],
     soln:solve(eq,A[i][j]), if length(soln) = 2 then
     solns:append(solns,ratsimp([soln[2]])) else
     solns:append(solns,ratsimp(soln)) )$
(C9) map(f90,ratsimp(solns))$
```

```
aa(1,1) = SQRT(cc(1,1))
aa(2,1) = cc(2,1)/aa(1,1)
aa(2,2) = SQRT(cc(2,2)-aa(2,1)**2)
aa(3,1) = cc(3,1)/aa(1,1)
aa(3,2) = (cc(3,2)-aa(2,1)*aa(3,1))/aa(2,2)
aa(3,3) = SQRT(cc(3,3)-aa(3,2)**2-aa(3,1)**2)
aa(4,1) = cc(4,1)/aa(1,1)
aa(4,2) = (cc(4,2)-aa(2,1)*aa(4,1))/aa(2,2)
aa(4,3) = (cc(4,3)-aa(3,2)*aa(4,2)-aa(3,1)*aa(4,1))/aa(3,3)
aa(4,4) = SQRT(cc(4,4)-aa(4,3)**2-aa(4,2)**2-aa(4,1)**2)
aa(5,1) = cc(5,1)/aa(1,1)
aa(5,2) = (cc(5,2)-aa(2,1)*aa(5,1))/aa(2,2)
aa(5,3) = (cc(5,3)-aa(3,2)*aa(5,2)-aa(3,1)*aa(5,1))/aa(3,3)
aa(5,4) = (cc(5,4)-aa(4,3)*aa(5,3)-aa(4,2)*aa(5,2)-aa(4,1)* &
     aa(5,1))/aa(4,4)
aa(5,5) = SQRT(cc(5,5)-aa(5,4)**2-aa(5,3)**2-aa(5,2)**2-aa(5,1)** &
     2)
aa(6,1) = cc(6,1)/aa(1,1)
aa(6,2) = (cc(6,2)-aa(2,1)*aa(6,1))/aa(2,2)
aa(6,3) = (cc(6,3)-aa(3,2)*aa(6,2)-aa(3,1)*aa(6,1))/aa(3,3)
aa(6,4) = (cc(6,4)-aa(4,3)*aa(6,3)-aa(4,2)*aa(6,2)-aa(4,1)* &
     aa(6,1))/aa(4,4)
aa(6,5) = (cc(6,5)-aa(5,4)*aa(6,4)-aa(5,3)*aa(6,3)-aa(5,2)* &
     aa(6,2)-aa(5,1)*aa(6,1))/aa(5,5)
aa(6,6) = SQRT(cc(6,6)-aa(6,5)**2-aa(6,4)**2-aa(6,3)**2-aa(6,2)** &
     2-aa(6,1)**2)
```

(C10)

Finite-statistics case

We now consider of (finite) set of n vectors $\{\rho^i\}$ whose components are (nominally) uncorrelated random numbers evenly distributed between 0 and 1. These vectors will have mean

$$\bar{\rho} \equiv \langle \rho_i \rangle.$$

and covariance matrix

$$X \equiv \langle \rho_i \rho_j \rangle.$$

As $n \to \infty$, $\bar{\rho} \to 0$ and $X \to \mathbb{1}$. This time, we seek a transformation of the form

$$r = A(\rho - \rho_0),$$

which leads us to the equations

$$\rho_0 = \bar{\rho},$$

and

$$C = \mathrm{AXA}^T.$$

In order to solve for A, we perform two Cholesky decompositions:

$$C = \mathrm{GG}^T$$

and

$$X = \mathrm{HH}^T.$$

The solution for A is[1]

$$A = \mathrm{GH}^{-1},$$

as can be easily verified by substitution. Practically, it is easy to calculate A by solving the equation

$$\mathrm{AH} = G.$$

```
(C10) A:genmatrix(aa,dim,dim)$
(C11) H:genmatrix(hh,dim,dim)$
(C12) G:genmatrix(gg,dim,dim)$
(C19) for i:1 thru dim do for j:i+1 thru dim do block( A[i][j]:0, H[i][j]:0,
      G[i][j]:0)$
(C20) AH:A.H$
(C21) solns:[]$
(C22) for i:1 thru dim do for j:i step -1 thru 1 do block( eq:AH[i][j] = G[i][j],
      soln:solve(eq,A[i][j]), solns:append(solns,ratsimp(soln)) )$
(C23) map(f90,ratsimp(solns))$
```

```
aa(1,1) = gg(1,1)/hh(1,1)
aa(2,2) = gg(2,2)/hh(2,2)
aa(2,1) = -(hh(2,1)*aa(2,2)-gg(2,1))/hh(1,1)
aa(3,3) = gg(3,3)/hh(3,3)
aa(3,2) = -(hh(3,2)*aa(3,3)-gg(3,2))/hh(2,2)
aa(3,1) = -(hh(3,1)*aa(3,3)+hh(2,1)*aa(3,2)-gg(3,1))/hh(1,1)
aa(4,4) = gg(4,4)/hh(4,4)
aa(4,3) = -(hh(4,3)*aa(4,4)-gg(4,3))/hh(3,3)
aa(4,2) = -(hh(4,2)*aa(4,4)+hh(3,2)*aa(4,3)-gg(4,2))/hh(2,2)
aa(4,1) = -(hh(4,1)*aa(4,4)+hh(3,1)*aa(4,3)+hh(2,1)* &
     aa(4,2)-gg(4,1))/hh(1,1)
aa(5,5) = gg(5,5)/hh(5,5)
aa(5,4) = -(hh(5,4)*aa(5,5)-gg(5,4))/hh(4,4)
aa(5,3) = -(hh(5,3)*aa(5,5)+hh(4,3)*aa(5,4)-gg(5,3))/hh(3,3)
aa(5,2) = -(hh(5,2)*aa(5,5)+hh(4,2)*aa(5,4)+hh(3,2)* &
     aa(5,3)-gg(5,2))/hh(2,2)
aa(5,1) = -(hh(5,1)*aa(5,5)+hh(4,1)*aa(5,4)+hh(3,1)* &
     aa(5,3)+hh(2,1)*aa(5,2)-gg(5,1))/hh(1,1)
aa(6,6) = gg(6,6)/hh(6,6)
aa(6,5) = -(hh(6,5)*aa(6,6)-gg(6,5))/hh(5,5)
aa(6,4) = -(hh(6,4)*aa(6,6)+hh(5,4)*aa(6,5)-gg(6,4))/hh(4,4)
aa(6,3) = -(hh(6,3)*aa(6,6)+hh(5,3)*aa(6,5)+hh(4,3)* &
     aa(6,4)-gg(6,3))/hh(3,3)
aa(6,2) = -(hh(6,2)*aa(6,6)+hh(5,2)*aa(6,5)+hh(4,2)* &
     aa(6,4)+hh(3,2)*aa(6,3)-gg(6,2))/hh(2,2)
aa(6,1) = -(hh(6,1)*aa(6,6)+hh(5,1)*aa(6,5)+hh(4,1)* &
     aa(6,4)+hh(3,1)*aa(6,3)+hh(2,1)*aa(6,2)-gg(6,1))/hh(1,1)
```

```
(C24)
```

1. Thanks to Mark Fischler for this solution.