Corrigendum: Behavior of the NORTA Method for Correlated Random Vector Generation as the Dimension Increases

SOUMYADIP GHOSH
IBM Research

and

SHANE G. HENDERSON
Cornell University

This note corrects an error in Ghosh and Henderson [2003].

Categories and Subject Descriptors: G.3 [Mathematics of Computing]: Probability and Statistics—Random number generation

General Terms: Theory, Algorithms

Additional Key Words and Phrases: NORTA method, semidefinite programming, sampling random matrices, onion method

ACM Reference Format:
DOI = 10.1145/1596519.1596525 http://doi.acm.org/10.1145/1596519.1596525

As part of their analysis of the NORTA method, Ghosh and Henderson [2003] developed an algorithm for generating a random correlation matrix. The derivation of the algorithm is correct up until the discussion on page 288 of the sampling scheme for the completion vector \( q \). Specifically, in the last step of the sampling scheme where \( q \) is set to \( \Sigma_{k-1}^{1/2} w \), the matrix \( \Sigma_{k-1}^{1/2} \) is not explicitly defined in the text. This corrigendum gives the definition that is consistent with the original description of the algorithm, and also presents an alternate definition that, with a slightly modified step, gives a computationally more efficient sampling scheme.

Authors’ addresses: S. Ghosh, IBM T. J. Watson Research Center, 1101 Kitchawan Rd, Yorktown Heights, NY 10598; email: ghosh@us.ibm.com; S. G. Henderson, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

The transformation \( w = \Sigma_{k-1}^{-1/2} q \) is introduced on page 287 in the discussion after Equation (14), where \( \Sigma_{k-1}^{-1/2} \) is defined to be the upper-triangular Cholesky factor of \( \Sigma_{k-1}^{-1} \). Consistent with this definition, the correct definition of \( \Sigma_{k-1}^{-1/2} \) as it appears in Ghosh and Henderson [2003] is that it is the inverse of the Cholesky factor \( \Sigma_{k-1}^{-1} \). The notation suggests, however, that it is the upper-triangular Cholesky factor of \( \Sigma_{k-1} \), which is incorrect since in general the two matrices are different.

With the correct definition, the computational burden of the last step of the method can be high: for each iteration that adds an extra layer, the principal submatrix \( \Sigma_{k-1}^{-1/2} \) constructed in the earlier pass must be inverted, its Cholesky factor computed, and then this factor inverted again. While incremental update schemes can be used to reduce the effort, it would be better if the transformation from \( w \) to \( q \) could be carried out using a factorization of the matrix \( \Sigma_{k-1}^{-1/2} \) itself.

The transformation \( w = \Sigma_{k-1}^{-1/2} q \) on page 287 requires only that \( \Sigma_{k-1}^{-1/2} \) be any matrix \( A \) satisfying \( A^T A = \Sigma_{k-1}^{-1} \), and not necessarily the upper-triangular Cholesky factor. Accordingly, we recommend taking \( A = (C')^{-1} \), where the matrix \( C \) is the upper-triangular Cholesky factor of \( \Sigma_{k-1} \). To see why this works, notice that \( C^T C = \Sigma_{k-1} \), and so \( C^{-1}(C')^{-1} = \Sigma_{k-1} \). So we may take \( w = (C')^{-1}q \) and hence \( q = Cw \). Notice that \( C \) can be constructed iteratively as the algorithm proceeds, leading to computational efficiency.

To summarize, redefining the matrix \( \Sigma_{k-1}^{-1/2} \) as the upper-triangular Cholesky factor of \( \Sigma_{k-1} \), the sampling scheme for generating the completion vector \( q \) is as follows.

1. Sample \( y \) from a beta distribution with \( \alpha_1 = (k-1)/2 \) and \( \alpha_2 = (d-k)/2 + 1 \).
2. Set \( r = \sqrt{y} \).
3. Sample a unit vector \( \theta \) uniformly from the surface of \( B_{k-1} \).
4. Set \( w = r\theta \), and finally
5. Set \( q = (\Sigma_{k-1}^{-1/2})^T w \).

This description also corrects a typographical error in stating the value of the parameter \( \alpha_2 \) in the original article, as previously reported in Ghosh and Henderson [2005].

The errors were confined to the description of the algorithm. The code used to generate the results in Ghosh and Henderson [2003] implemented the modified scheme presented above.

ACKNOWLEDGMENTS

We would like to thank Seksan Kiatsupaibul in helping us bring this error to light.

REFERENCES

Corrigendum: Behavior of the NORTA Method


Received June 2009; accepted July 2009