Hierarchical-block conditioning approximations for high-dimensional multivariate normal probabilities

- **Hyunseok Yang (@YangHS329)**
- Kyeongwon Lee (@kw-lee)
- Songhyun Kim (@than0you)

The name of the leader is written in bold. All members are affiliated with Departments of Statistics, Seoul National University.

Brief Introduction

The computation of multivariate normal probability appears various fields. For instance, the inferences based on the central limit theorem, which holds when the sample size is large enough, is widely used in the social sciences and engineering as well as in the natural sciences. Recently, the dimensionality of data and models has been grown significantly, and in this respect, so does a need for the methodology to efficiently calculate high-dimensional multivariate normal probability.

Cao, et al. (2019)[1] proposes new approaches to approximate high-dimensional multivariate normal probability

$$\Phi_n(a, b; 0, \Sigma) = \int_a^b \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\right) d\mathbf{x}$$

using the hierarchical matrix $\mathcal{H}$ (Hackbusch (2015)[2]) for the covariance matrix Σ. The methods are based on two state-of-arts methods, among others, are the bivariate conditioning method (Trinh and Genz (2015)[3]) and the hierarchical Quasi-Monte Carlo method (Genton et al. (2018)[4]). Specifically, Cao et al. (2019) generalize the bivariate conditioning method to a d-dimension and combine it with the hierarchical representation of the covariance matrix.

Goal

The main goal of the project is to find **good** approximations for high-dimensional multivariate normal probabilities. Specifically,

- Review the methods proposed by Cao et al. (2019).
- Compare to existing methods including what was covered in lecture.
- Reproduce the results in Cao et al. (2019).

Plan

This project has a five-point plan:

1. Precedent research
   The author mentioned that the computation method adopted by this paper is based on the following two methods: (i) Bivariate conditioning method, and (ii) hierarchical Quasi-Monte Carlo method, respectively introduced by Trin and Genz (2015) and Genton et al. (2018). Therefore, our group firstly will study the two precedent procedures for computing MVN pdf. Also, the effect of reordering when it comes to the derivation of computing truncated moments should also be in consideration.
2. Review the paper
   The paper built a hierarchical covariance matrix based on the method of hierarchical representation introduced by Hackbusch (2015). When off-diagonal locations have low-rank feature while diagonal blocks are dense, the covariance matrix can be arranged with hierarchical representation. Since the paper does not contain detailed knowledge regarding to the theory, our group will further investigate about the overall algorithm of generating hierarchical representation.
3. Implementation with Julia
   The authors introduced the computation results by calculating exponential covariance model using three methods: (i) Cholesky factorization by R (ii) Cholesky factorization by `LAPACK` (iii) hierarchical Cholesky factorization by `H2Lib` . We are planning to reproduce the computational analysis introduced above with various covariance model and a number of parameters. Primary programming tool would be Julia, but other tools such as R can be used if they are needed.
4. Compare to existing methods
   Our group is aiming to search for theoretical explanation about the methods that was tersely introduced. Especially, we are looking for the reason why Quasi-Monte-Carlo method is adopted. This would be done by studying the paper written by Genton et al. (2018). Also, our group will try to figure out the specific reasoning that reordering of parameters enhances the performance of computation. For this, we study the paper of Trinh and Genz (2015).
5. Apply to real data
   If time allows, we will apply these methods to real data. We plan to use spatial data as in the paper.

References

1. Cao, J., Genton, M. G., Keyes, D. E., & Turkiyyah, G. M. (2019). Hierarchical-block conditioning approximations for high-dimensional multivariate normal probabilities. Statistics and Computing, 29(3), 585-598.↩
2. Hackbusch, W. (2015). Hierarchical matrices: algorithms and analysis (Vol. 49). Heidelberg: Springer.↩
3. Trinh, G., & Genz, A. (2015). Bivariate conditioning approximations for multivariate normal probabilities. Statistics and Computing, 25(5), 989-996.↩
4. Genton, M. G., Keyes, D. E., & Turkiyyah, G. (2018). Hierarchical decompositions for the computation of high-dimensional multivariate normal probabilities. Journal of Computational and Graphical Statistics, 27(2), 268-277.↩