Correlation Structures for Optimizing Information Criteria

Aleksandar Kavčić
Harvard University
Division of Engineering and Applied Sciences
29 Oxford Street, Cambridge, MA 02138
Email: kavcic@hrl.harvard.edu

José M. F. Moura
Carnegie Mellon University
Dept. of Electrical and Computer Engineering
5000 Forbes Avenue, Pittsburgh, PA 15213
Email: moura@ece.cmu.edu

Abstract — **We consider the problem of approximating general Gauss processes by a Gauss-Markov random process (GMrp). GMrp's have covariances whose inverses are banded. The optimality criteria are maximum entropy and minimum cross-entropy (Kullback-Leibler distance). We apply to this problem a new banded matrix inversion theorem and a new matrix decomposition theorem. Our solution is an alternative to solving Yule-Walker equations or Cholesky decompositions. Its special structure offers advantages in Viterbi sequence detection and provides closed form expressions to the inverses of certain general banded matrices.**

I. Banded matrix algebra

Notation An L-banded matrix has zeros above the L-th upper and below the L-th lower diagonals. The L-band of a matrix is the band contained between the L-th upper and lower diagonals. The matrix $\mathbf{R}_i^j$ denotes the principal minor of the matrix $\mathbf{R}$ defined by columns i through j. The proofs of the following theorems are in [1].

Theorem 1 (Banded Matrix Inverse Theorem) *Let $\mathbf{R}$ be a matrix whose inverse $\mathbf{R}^{-1}$ is L-banded. The inverse is then given by*

$$\mathbf{R}^{-1} = \left[\begin{array}{c} \boxed{\mathbf{R}_{L+1}^{1}}^{-1} - \boxed{\mathbf{R}_{L+1}^{2}}^{-1} + \boxed{\mathbf{R}_{L+2}^{2}}^{-1} - \cdots - \boxed{\mathbf{R}_{N-1}^{N-L}}^{-1} + \boxed{\mathbf{R}_{N}^{N-L}}^{-1} \end{array} \right]$$

This notation denotes that each principal minor is inverted and added to its corresponding position.

Theorem 2 (Decomposition Theorem) *Let $\mathbf{C}$ be an arbitrary square matrix. There exists a unique matrix $\mathbf{R}$ whose L-band equals the L-band of $\mathbf{C}$ and whose inverse $\mathbf{R}^{-1}$ is L-banded. In other words, there exists a unique matrix $\mathbf{R}$ such that*

$$\mathbf{C} = \mathbf{R} + \begin{bmatrix} & \boldsymbol{\Phi} \\ \mathbf{0} & \\ \boldsymbol{\Psi} & \end{bmatrix} \quad \text{and} \quad \mathbf{R}^{-1} = \begin{bmatrix} & \mathbf{0} \\ \mathbf{D} & \\ \mathbf{0} & \end{bmatrix},$$

where $\mathbf{R}^{-1}$ is L-banded.

Definition 1 (L-band Complement) *The matrix $\mathbf{R}$ in (2) is the L-band complement of $\mathbf{C}$.*

II. Optimal correlation structures

The proofs of the followig theorems are in [1].

Theorem 3 (Kullback-Leibler Optimal Covariance) *Let $\mathbf{C}$ be the covariance matrix of a zero-mean Gaussian vector $\underline{z}$. Let $\mathbf{R}$ be the covariance matrix of the L-th order Gauss-Markov approximation vector $\dot{\underline{z}}$. The matrix $\mathbf{R}$ that minimizes the cross-entropy (Kullback-Leibler distance) between $\underline{z}$ and $\dot{\underline{z}}$ is the L-band complement of $\mathbf{C}$.*

Theorem 4 (Maximum Entropy Covariance) *Let $\mathbf{C}$ be a non-Toeplitz covariance matrix. Let $\dot{\underline{z}}$ be a zero-mean random vector, such that the L-band of its covariance matrix $\mathbf{R}$ equals the L-band of $\mathbf{C}$. The entropy of $\dot{\underline{z}}$ is maximized when $\dot{\underline{z}}$ is an L-th order Gauss-Markov vector whose covariance matrix $\mathbf{R}$ is the L-band complement of $\mathbf{C}$.*

III. Discussion

Theorems 3 and 4 establish that the L-band complement provides the optimal L-th order Gauss-Markov process approximation. Our solution shows the equivalence between the L-band complement and the solution obtained by solving for the corresponding autoregressive parameters through general nonstationary Yule-Walker equations, [2, 3]. While the two approaches are equivalent, choosing one over the other may prove beneficial in different cases. The L-band approach, for example, leads to elegant applications and interpretations of the Viterbi algorithm in the presence of correlated noise [4] as well as to closed-form solutions to inverting arbitrary tridiagonal matrices [1].

References

[1] A. Kavčić and J. M. F. Moura, "Gauss-Markov approximations and information loss," tech. rep., Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, March 1998. 30 pages, submitted for publication.

[2] J. P. Burg, *Maximum Entropy Spectral Analysis.* PhD thesis, Department of Geophysics, Stanford University, Stanford, CA, 1975.

[3] H. Lev-Ari, S. R. Parker, and T. Kailath, "Multidimensional maximum-entropy covariance extension," *IEEE Trans. Inform. Theory*, vol. 35, pp. 497–508, May 1989.

[4] A. Kavčić and J. M. F. Moura, "The Viterbi algorithm and Markovian noise memory," tech. rep., Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, November 1997. 30 pages, submitted for publication.