A regularized Cepstrum and Covariance Matching method for ARMA(n,m) design

Per Enqvist
LADSEB/CNR
Corso Stati Uniti 4
35 127 Padova
Italy

Abstract

An ARMA(n,m) model is uniquely determined by its n first covariances and m first cepstrum coefficients. However, there does not always exist a model matching an estimated set of these parameters. We propose a method determining an asymptotically stable minimum phase model that match the covariances exactly and the cepstrum parameters approximatively. A convex barrier term is used for the regularization and the model is determined by a convex optimization problem.

1 Introduction

The maximum entropy method has been used to solve the covariance interpolation problem for models of stochastic processes. Using this approach, an all-pole rational model, often called just the Maximum Entropy (ME) model, is obtained. Generalizing the ME approach to also use information about the zeros of the model, a pole-zero rational model is obtained [1, 2]. Instead of using direct information about the zeros, it was shown in [2, 3] that using information on the cepstrum parameters of the model, in addition to the covariances, also determines a unique pole-zero rational model.

It should be noted that applying the maximum entropy approach to any arbitrary combination of observables, a complex non-linear model is generally obtained, but in this particular case a finite order rational model maximizes the entropy. In fact, as in the ME problem, the basic maximum entropy problem is an infinite dimensional problem that is solved by considering a dual problem. The existence of an exactly matching model is connected to the solution of the corresponding dual optimization problem beeing an interior point. For the ME problem and the problem with fixed zeros, there is always an interior point solution, but for the covariance and cepstrum problem, there might be solutions on the boundary. This means that for some combinations of covariances and cepstrum parameters there might not exist an interpolating model, which can be compared to the problem with fixed zeros, where for any zero-polynomial there always exist a rational model interpolating the covariances. Anyway, since the cepstrum parameters can be estimated directly from data, it is tractable to consider the covariance and cepstrum interpolation problem.

However, since there does not always exist an ARMA(n,m) model matching a generic set of data exactly, some approximation has to be done. In order to obtain interior point solutions for generic data, a barrier like term times a weighting parameter is added to the objective function of the dual optimization problem. It is shown that the introduction of the barrier term forces the solution to the interior, it increases the entropy of the solution, and further that the entropy increases monotonically with the weighting parameter of the barrier term. If the weighting parameter tend to infinity, the solution will tend to the maximum entropy solution corresponding to the covariance data, and if the weighting term is set to zero the original cepstrum and covariance matching problem is obtained. Since the barrier term is also convex, the proposed optimization problem is convex and has a unique solution in an interior point corresponding to an ARMA(n,m) model with poles and zeros strictly inside the unit circle. Moreover, since the map from data to model is analytic, a small change in the data will result in a small change in the resulting model, which provides robustness of the proposed regularized CCM method.

2 Preliminaries

The window of covariance data $\mathcal{R}_n = \{r_0, r_1, \cdots, r_n\}$ is assumed to correspond to a *bona fide* covariance sequence, which means that

$$\mathbf{R}_n \triangleq \begin{bmatrix} \hat{r}_0 & \hat{r}_1 & \dots & \hat{r}_n \\ \hat{r}_1 & \hat{r}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \hat{r}_1 \\ \hat{r}_n & \dots & \hat{r}_1 & \hat{r}_0 \end{bmatrix} \tag{2.1}$$

is positive definite. The window of cepstrum data $\mathcal{C}_m = \{c_1, \cdots, c_m\}$ is the Fourier coefficients of the logarithm of the spectral density Φ,

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\log\Phi(e^{i\theta})d\theta, \quad k = 0, 1, 2, \dots, \tag{2.2}$$

where the spectral density is defined as

$$\Phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} r_k(e^{-ik\theta}). \tag{2.3}$$

A finite window of the cepstrum coefficients can be estimated directly from data using ergodic estimates. In particular, taking the DFT of the data, the absolute value square, the logarithm and the IDFT determines an estimate of the cepstrum [4, 5].

The transfer function of a stable and minimum phase ARMA(n,m) model can be described by

$$W(z) = \frac{\sigma(z)}{a(z)},$$

where

$$a(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n, \qquad (a_0 > 0), \tag{2.4}$$

$$\sigma(z) = z^m + \sigma_1 z^{m-1} + \cdots + \sigma_m \tag{2.5}$$

are polynomials having all their roots in the open unit disc. If W is the transfer function of an ARMA(n,m) model whose input is variance one white noise, then $\Phi(z) = W(z)W(z^{-1})$ is the corresponding spectral density. Then

$$\Phi = P/Q,$$

where P and Q are pseudo polynomials

$$P(z) = p_0 + \frac{1}{2}p_1(z + z^{-1}) + \cdots + \frac{1}{2}p_m(z^m + z^{-m}), \tag{2.6}$$

$$Q(z) = q_0 + \frac{1}{2}q_1(z + z^{-1}) + \cdots + \frac{1}{2}q_n(z^n + z^{-n}), \tag{2.7}$$

where, since $P(z) = \sigma(z)\sigma(z^{-1})$ and $Q(z) = a(z)a(z^{-1})$, both P and Q are positive on the unit circle. Given such a pseudo-polynomial P, using spectral factorization it is possible to determine the unique polynomial with roots in the unit disc such that $\sigma(z)\sigma(z^{-1}) = P(z)$. We denote the set of such positive pseudo-polynomials by $\mathcal{D}_n$, where n denotes the degree. A stable and minimum phase ARMA(n,m) model can then be described by two pseudo-polynomials $P \in \mathcal{D}_m$ and $Q \in \mathcal{D}_n$.

The dual optimization problem characterizing the maximum entropy solution to the the covariance and cepstrum interpolation problem [2], can be formulated as

$$(\mathcal{P}) \quad \begin{bmatrix} \min & \varphi(P, Q), \\ \text{s.t.} & Q \in \mathcal{D}_n, P \in \mathcal{D}_m, p_0 = 1. \end{bmatrix}$$

where the objective function is given by

$$\varphi(P, Q) \triangleq -\sum_{k=1}^{m} c_k p_k + \sum_{k=0}^{n} r_k q_k + \frac{1}{2\pi}\int_{-\pi}^{\pi} P \log\frac{P}{Q} d\theta. \tag{2.8}$$

3 Regularization of Problem ($\mathcal{P}$)

Since the optimization problem ($\mathcal{P}$) may have solutions on the boundary of the feasible region, a regularization of that problem is proposed in this section. In order to force the solution to the interior of the feasible region, the optimization problem ($\mathcal{P}$) is modified by adding a barrier like term to the objective function φ. By adding the term $\beta = -\langle 1, \log P\rangle$ to the objective function φ of the problem ($\mathcal{P}$) the optimal P is forced into the interior of $\mathcal{D}_m$.

It follows from [1] that $\beta = -\langle 1, \log P\rangle$ has bounded level sets, is finite for all $P \in \mathcal{D}_m$ such that $p_0 = 1$ and has a derivative that tends to infinity as P tends to the boundary of $\mathcal{D}_m$. It is further strictly convex in P. Since the function values of β are finite at the boundary, it is not a barrier function from a mathematical programming view point [6], but since the derivatives of φ with respect to P are infinite, it will constrain the optimum away from the boundary as a barrier function would. It is shown later that this term will also increase the entropy of the solution.

The regularized optimization problem is formulated as

$$(\mathcal{P}_\lambda) \quad \left[\begin{array}{ll} \min & \varphi_\lambda(P,Q), \\ \text{s.t.} & Q \in \mathcal{D}_n, P \in \mathcal{D}_m, p_0 = 1, \end{array} \right]$$

where the modified objective function is given by

$$\varphi_\lambda(P,Q) \triangleq -\sum_{k=1}^{m} c_k p_k + \sum_{k=0}^{n} r_k q_k + \frac{1}{2\pi}\int_{-\pi}^{\pi} P \log\frac{P}{Q} d\theta - \frac{1}{2\pi}\int_{-\pi}^{\pi} \lambda \log P d\theta. \tag{3.9}$$

It is easy to see that φ_λ is strictly convex on a convex domain. If the minimum is located in the interior, the gradient

$$\frac{\partial \varphi_\lambda}{\partial p_k} = -c_k + \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta} \log\frac{P}{Q} d\theta - \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{\lambda}{P} d\theta, \qquad k = 1, \ldots, m, \tag{3.10}$$

$$\frac{\partial \varphi_\lambda}{\partial q_k} = r_k - \frac{1}{2\pi}\int_{-\pi}^{\pi} z^k \frac{P}{Q} d\theta, \qquad k = 0, 1, \ldots, n, \tag{3.11}$$

is zero at this point. The barrier term will make sure that the optimum will occur at an interior point. Hence, we have the following theorem.

Theorem 3.1. *For each $\lambda > 0$, Problem $(\mathcal{P}_\lambda)$ has a unique solution in the open set $\mathcal{D}_m \times \mathcal{D}_n$, At this point*

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{P}{Q} d\theta &= r_k, \qquad && k = 0, 1, \ldots, n \qquad (3.12) \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \langle e^{ik\theta}, \log\frac{P}{Q}\rangle d\theta &= c_k + \lambda\epsilon_k, && k = 1, \ldots, m, \qquad (3.13) \end{aligned}$$

where $\epsilon_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\theta}\frac{1}{P} d\theta$.

For the proof see [7, 8].

3.1 The Barrier Term and Entropy

In practice, the weight λ should be chosen smaller the better the data is, and the larger the model order is relative to the complexity of the system to be realized.

The entropy of a process is defined as

$$\mathcal{E}(\Phi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log \Phi d\theta, \tag{3.14}$$

see for example [9]. If the regularization term would have been chosen as $-\lambda\mathcal{E}(\Phi)$, the new problem would not have been convex and the resulting model would not have interpolated the covariances exactly. The choice of the regularization term proposed here is equivalent to this entropy term for fixed Q, and it can be shown that the resulting model will have an increased entropy even for this choice.

Proposition 3.1. *The entropy of the filter Φ_λ, that solves the problem $(\mathcal{P}_\lambda)$, is a monotonic non-decreasing function of λ.*

For the proof see [7, 8]. The barrier term is thus a measure of the entropy of the process. Then λ is a weight that determines how much the entropy should be maximized relative to the importance of the interpolation.

Standard optimization algorithms can be used to solve $(\mathcal{P}_\lambda)$. For example a damped Newton method, similar to the one proposed in [2] for the problem with fixed zeros, can be used. The best performance is probably achieved by using a homotopy method similar to the one presented in [10]. For numerical experiments see [7, 8].

References

[1] C.I. Byrnes, S.V. Gusev, and A. Lindquist. A convex optimization approach to the rational covariance extension problem. *SIAM Journal on Control and Optimization*, 37(1):211–229, 1999.

[2] C.I. Byrnes, P. Enqvist, and A. Lindquist. Cepstral coefficients, covariance lags and pole-zero models for finite data strings. *IEEE Trans. Signal Processing*, pages 677–693, April 2001.

[3] C.I. Byrnes, P. Enqvist, and A. Lindquist. Identifiability and well-posedness of shaping-filter parameterizations: a global analysis approach. *SIAM J. on Control and Optimization*, 41(1):23–59, 2002.

[4] L.R. Rabiner and R.W. Schafer. *Digital Processing of Speech Signals.* Prentice Hall, 1978.

[5] Y. Ephraim and M. Rahim. On second-order statistics and linear estimation of cepstral coefficients. *IEEE Trans. on Speech and Audio Processing*, 7(2):162–176, March 1999.

[6] D.G. Luenberger. *Linear and Nonlinear Programming.* Addison Wesley, 1984.

[7] P. Enqvist. *Spectral Estimation by Geometric, Topological and Optimization Methods.* PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2001.

[8] P. Enqvist. A convex optimization approach to ARMA(n,m) model design from covariance and cepstrum data. Submitted for publication to SIAM Journal on Control and Optimization.

[9] J.P. Burg. *Maximum Entropy Spectral Analysis.* PhD thesis, Stanford University, 1975.

[10] P. Enqvist. A homotopy approach to rational covariance extension with degree constraint. *International Journal of Applied Mathematics and Computer Science*, 11(5):1173–1201, 2001.