Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 00–39

Subspace angles and distances between ARMA models[1]

Katrien De Cock and Bart De Moor[2]

March 2000
Published in Proceedings of the Fourteenth International Symposium of Mathematical Theory of Networks and Systems (MTNS 2000)
Perpignan, France, June 19–23, 2000

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/decock/reports/00–39.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 17 09, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *katrien.decock@esat.kuleuven.ac.be, bart.demoor@esat.kuleuven.ac.be*. Bart De Moor is a Research Associate with the F.W.O. (Fund for Scientific Research-Flanders) and professor extra-ordinary at the K.U.Leuven. Katrien De Cock is a Research Assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). This work is supported by several institutions: the Flemish Government (Research Council K.U.Leuven: Concerted Research Action Mefisto-666, the FWO projects G.0240.99, G.0256.97, and Research Communities: ICCoS and ANMMM, IWT projects: EUREKA 1562-SINOPSYS, EUREKA 2063-IMPACT, STWW), the Belgian State, Prime Minister's Office (IUAP P4-02 (1997-2001) and IUAP P4-24 (1997-2001), Sustainable Mobility Program - Project MD/01/24 (1997-2000)), the European Commission (TMR Networks: ALAPEDES and System Identification, Brite/Euram Thematic Network: NICONET), Industrial Contract Research (ISMC, Data4S Electrabel, Laborelec, Verhaert, Europay)

Subspace angles and distances between ARMA models

Katrien De Cock, Bart De Moor
Katholieke Universiteit Leuven
Dept. of Electrical Engineering (ESAT), Research group SISTA
Kardinaal Mercierlaan 94, B-3001 Leuven, Belgium
Tel. 32/16 32 17 09, Fax 32/16 32 19 70
http://www.esat.kuleuven.ac.be/sista
{katrien.decock,bart.demoor}@esat.kuleuven.ac.be

Keywords: principal angles, ARMA models, linear systems, stochastic realization, distance measure

Abstract

In this paper we define a notion of subspace angles between two ARMA models as the principal angles between the column spaces generated by the observability matrices of the two models extended with the observability matrices of the inverse models. We show that these angles can be calculated via the observability Gramians of concatenations of the two given models and their inverses. Furthermore, we give their relation to different distance measures for ARMA models. Most attention is paid to a recently defined metric for ARMA models that is based on the cepstra of the models involved.

1 Introduction

The concept of principal angles between subspaces of linear vector spaces is due to Jordan [11] in the previous century. This notion was translated into the statistical notion of canonical correlations by Hotelling [10]. Applications include data analysis [6], random processes [5] [12] and stochastic realization [1] [3] [15] (and references herein). Numerically stable methods to compute the canonical structure via a singular value decomposition have been proposed by Björk and Golub [2], Golub and Zha [9] and can also be found in [8].
In this paper we define a notion of principal angles and their corresponding principal directions between two autoregressive-moving-average (ARMA) models and show how different distance measures between the models are related to the angles between them. The relation to the metric for ARMA models defined by Martin [14] in particular, will be discussed in more detail.
The models we consider are SISO–LTI (single input single output linear time invariant).
The paper is organized as follows. In Section 2 we briefly recall the definition of principal angles between and corresponding principal directions in two subspaces and we show how they can be computed. In Section 3 we discuss a distance measure for AR models, which has recently been defined by Martin [14]. Our definition of the subspace angles between AR models is given in Section 4, where we also give two methods for the computation of the angles. The relation between the subspace angles between two AR models and their distance as defined in [14] is given in Section 5. Two other distance measures that can be derived from the subspace angles are given in Section 6. The definition of the distances and subspace angles is extended to the ARMA model class in Section 7, where we also explain how to compute the subspace angles between two ARMA models. Finally in Section 8 we give the conclusions and point out possible further developments of our work.

2 Principal angles between subspaces

In this section we discuss the notion of principal angles between and principal directions in two subspaces. We start with the definition in Section 2.1. In Section 2.2 we show how the angles and directions can be characterized from a generalized eigenvalue problem and in Section 2.3 we give a method for the computation of the principal angles and directions via the singular value decomposition.

2.1 Definition

Let $A \in \mathbb{R}^{m\times p}$ and $B \in \mathbb{R}^{m\times q}$ be given real matrices with the same number of rows and assume for convenience that A and B have full column rank and that $p \geq q$. We denote the range (column space) of a matrix A by $\text{range}(A)$.

Definition 2.1. The q principal angles $\theta_k \in \left[0, \frac{\pi}{2}\right]$, between $\text{range}(A)$ and $\text{range}(B)$ are recursively defined for $k = 1, 2, \ldots, q$ as

$$\begin{aligned}\cos\theta_1 &= \max_{\substack{x\in\mathbb{R}^p\\ y\in\mathbb{R}^q}} \frac{|x^T A^T B y|}{\|Ax\|_2\|By\|_2} \\ &= \frac{|x_1^T A^T B y_1|}{\|Ax_1\|_2\|By_1\|_2},\end{aligned}$$

$$\begin{aligned}\cos\theta_k &= \max_{\substack{x\in\mathbb{R}^p\\ y\in\mathbb{R}^q}} \frac{|x^TA^TBy|}{\|Ax\|_2\|By\|_2} \\ &= \frac{|x_k^TA^TBy_k|}{\|Ax_k\|_2\|By_k\|_2}\ ,\ \text{for } k=2,\ldots,q\end{aligned}$$

subject to $x_i^TA^TAx = 0$ and $y_i^TB^TBy = 0$,
for $i = 1, 2, \ldots, k-1$.

Note that the principal angles satisfy

$$0 \le \theta_1 \le \ldots \le \theta_q \le \frac{\pi}{2}\,.$$

The vectors $Ax_1, \ldots, Ax_q$ and $By_1, \ldots, By_q$ are called the principal directions of the pair of spaces. Following the notation in [15], the set of q principal angles between the ranges of the matrices A and B is denoted by $[A \lhd B]$.

2.2 The principal angles and directions as the solution of a generalized eigenvalue problem

It can be shown (see e.g. [9]) that the principal angles and the principal directions between range(A) and range(B) follow from the symmetric generalized eigenvalue problem:

$$\begin{pmatrix} 0 & A^TB \\ B^TA & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} A^TA & 0 \\ 0 & B^TB \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\ ,$$
$$\text{subject to } x^TA^TAx = 1 \text{ and } y^TB^TBy = 1. \tag{1}$$

The link between the formula in (1) and Definition 2.1 goes via the so-called variational characterization of the eigenvalue problem.
Assume that the $p+q$ (real) eigenvalues λ_i are sorted in descending order as $\lambda_1 \ge \ldots \ge \lambda_{p+q}$, then one can show that

$$\begin{aligned}&\lambda_1 = \cos\theta_1, \ldots, \lambda_q = \cos\theta_q \ge 0\ ,\\ &\lambda_{p+q} = -\cos\theta_1, \ldots, \lambda_{p+1} = -\cos\theta_q\ ,\\ &\lambda_{q+1} = \lambda_{q+2} = \cdots = \lambda_p = 0\ .\end{aligned} \tag{2}$$

The vectors Ax_j and By_j, for $j = 1, \ldots, q$ where x_j and y_j satisfy (1) with $\lambda = \lambda_j$, are the principal directions corresponding to the principal angle θ_j.
Note that when considering the principal angles between equidimensional subspaces ($p = q$), Equation (2) does not come into play.

2.3 Computing the principal angles using the singular value decomposition (SVD)

Let the columns of $U_1 \in \mathbb{R}^{m\times p}$ and $U_2 \in \mathbb{R}^{m\times q}$ generate orthonormal bases for range(A) and range(B) respectively, and let

$$U_1^TU_2 = U\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_q \end{pmatrix}V^T$$

be the SVD of $U_1^TU_2$, where $U \in \mathbb{R}^{p\times q}$ and $V \in \mathbb{R}^{q\times q}$. The singular values σ_i are then equal to the cosines of the principal angles between range(A) and range(B) and the columns of U_1U and U_2V are the principal directions (for a proof see e.g. [8]).
The orthonormal bases U_1 and U_2 can be obtained by computing the QR-decomposition of A and B:

$$\begin{aligned}A &= Q_1R_1\ ,\ \text{where } Q_1^TQ_1 = I_p\ ,\\ B &= Q_2R_2\ ,\ \text{where } Q_2^TQ_2 = I_q\ ,\end{aligned}$$

and taking $U_1 = Q_1$ and $U_2 = Q_2$, or by computing their SVD:

$$\begin{aligned}A &= U_1S_1V_1^T\ ,\ \text{where } U_1^TU_1 = I_p\ ,\\ B &= U_2S_2V_2^T\ ,\ \text{where } U_2^TU_2 = I_q\ .\end{aligned}$$

The SVD approach for the computation of the orthonormal basis of A and B has the advantage that it can be used when A and/or B does not have full column rank. However, computing the SVD of A and B requires much more work than computing the corresponding QR-decompositions [2].

3 A metric for the set of SISO–LTI AR models

In [14] Martin defines a new metric for the set of SISO–LTI ARMA models. It is based on the inner product of the cepstra of ARMA models. Further on in the paper we will show that this metric is related to the principal angles between specific subspaces derived from the ARMA models. For the sake of completeness, we repeat in this section some results which have been reported in [14].
However, in this section, we restrict our discussion to the AR model class. The extension to the ARMA models will be made in Section 7.

Let M_1 and M_2 be AR models with cepstrum coefficients $c_1(n)$ and $c_2(n)$, $n = 0, \pm1, \pm2, \ldots$.

Definition 3.1. [14] The distance d_M between M_1 and M_2 is defined as

$$d_M(M_1, M_2) = \sqrt{\sum_{n=0}^{\infty} n|c_1(n) - c_2(n)|^2}\,. \tag{3}$$

The index M stands for 'Martin' and is used to distinguish the metric from other distance measures, discussed in this paper.
Martin subsequently shows that for stable AR models M_1 with order n_1 and poles α_i and M_2 with order n_2 and poles β_i the following equality holds

$$d_M(M_1, M_2)^2 = \ln \frac{\displaystyle\prod_{i=1}^{n_1}\prod_{j=1}^{n_2}|1-\bar\alpha_i\beta_j|^2}{\displaystyle\prod_{i,j=1}^{n_1}(1-\bar\alpha_i\alpha_j)\prod_{i,j=1}^{n_2}(1-\bar\beta_i\beta_j)}\,. \tag{4}$$

This equality basically follows from an expression that relates the cepstrum coefficients of an AR model to its poles, which can be found in [13].

Observe that if M_1 and M_2 are two stable first order AR models, their distance equals

$$d_M(M_1, M_2)^2 = \ln \frac{(1-\alpha\beta)^2}{(1-\alpha^2)(1-\beta^2)} = \ln \frac{1}{\cos^2\theta} ,$$

where θ is the angle between the vectors

$$\begin{pmatrix} 1 \\ \alpha \\ \alpha^2 \\ \vdots \end{pmatrix} \in \mathbb{R}^\infty \text{ and } \begin{pmatrix} 1 \\ \beta \\ \beta^2 \\ \vdots \end{pmatrix} \in \mathbb{R}^\infty .$$

It will become apparent in Section 5 that for higher order models, the squared distance as defined by Martin [14] can be expressed as the logarithm of a product of $\frac{1}{\cos^2\theta_i}$ (see Theorem 1). The angles θ_i will be called the subspace angles between the AR models.

4 Subspace angles between AR models

In this section we start the discussion of our new concept of angles between models, by considering AR models. The definition of the subspace angles between two AR models is given in Section 4.1. In Section 4.2 we show how the angles can be computed. We give two methods for the computation. In Section 4.2.1 the angles are obtained via the observability Gramian of a concatenation of the two models. The second method goes via the poles of the AR models and is briefly described in Section 4.2.2.

4.1 Definition

Let two stable and observable nth order AR models M_1 and M_2 be characterized in state space terms by their system matrix A_1 and A_2 and output matrix C_1 and C_2 respectively. Their infinite observability matrix,

$$\begin{pmatrix} C_i \\ C_iA_i \\ C_iA_i^2 \\ \vdots \end{pmatrix} \in \mathbb{R}^{\infty\times n} ,$$

is denoted by $\mathcal{O}_\infty(M_i)$ for $i = 1, 2$.

Definition 4.1. We define the subspace angles between M_1 and M_2 as the principal angles between the ranges of their infinite observability matrices:

$$[M_1 \lhd M_2] = [\mathcal{O}_\infty(M_1) \lhd \mathcal{O}_\infty(M_2)] .$$

If two AR models do not have the same order, and assume the difference in order is equal to Δn, then we add Δn poles in zero to the model with the smallest order. Consequently, the number of subspace angles between two AR models of order n_1 and n_2 respectively is always equal to $\max(n_1, n_2)$.

4.2 Computation of the subspace angles between AR models

We now show how the subspace angles between two AR models can be computed. In Section 4.2.1 this is done via the observability Gramian of a concatenation of the two models. However, when the poles of the models are known, the computation of the angles can be computed in a more efficient way. This is briefly discussed in Section 4.2.2.

4.2.1 Computation via the observability Gramian

For simplicity we take $n_1 = n_2 = n$.
Let the AR models M_1 and M_2 be characterized in state space terms by their system matrix A_1 and A_2 and output matrix C_1 and C_2 respectively. Their infinite observability matrix is denoted by $\mathcal{O}_\infty(M_1)$ and $\mathcal{O}_\infty(M_2)$. Assume the two models are stable and observable.

The cosines of the subspace angles between M_1 and M_2 are equal to the largest n eigenvalues of $\begin{pmatrix} 0 & Q_{11}^{-1}Q_{12} \\ Q_{22}^{-1}Q_{21} & 0 \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$, where $Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$ is the unique solution of the Lyapunov equation

$$\begin{pmatrix} A_1^T & 0 \\ 0 & A_2^T \end{pmatrix} Q \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} - Q = - \begin{pmatrix} C_1^T \\ C_2^T \end{pmatrix} \begin{pmatrix} C_1 & C_2 \end{pmatrix} . \tag{5}$$

This can be seen as follows.
The subspace angles between M_1 and M_2 are equal to the principal angles between range $(\mathcal{O}_\infty(M_1))$ and range $(\mathcal{O}_\infty(M_2))$ (Definition 4.1). Consequently, the cosines of the subspace angles between M_1 and M_2 are equal to (see Section 2.2) the n largest generalized eigenvalues of $\begin{pmatrix} 0 & Q_{12} \\ Q_{21} & 0 \end{pmatrix}$ and $\begin{pmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{pmatrix}$, where

$$\begin{aligned} Q &= \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \in \mathbb{R}^{2n\times 2n} \\ &= \begin{pmatrix} \mathcal{O}_\infty(M_1)^T \\ \mathcal{O}_\infty(M_2)^T \end{pmatrix} \begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2) \end{pmatrix} . \end{aligned}$$

The matrix Q is the observability Gramian of the linear system with system matrix $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ and output matrix $\begin{pmatrix} C_1 & C_2 \end{pmatrix}$.
It is easy to show that Q satisfies the Lyapunov equation in (5).
Furthermore, since M_1 and M_2 are observable, the matrices Q_{11} and Q_{22} are non-singular. Therefore, the generalized eigenvalues of $\begin{pmatrix} 0 & Q_{12} \\ Q_{21} & 0 \end{pmatrix}$ and $\begin{pmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{pmatrix}$ are equal to the eigenvalues of $\begin{pmatrix} 0 & Q_{11}^{-1}Q_{12} \\ Q_{22}^{-1}Q_{21} & 0 \end{pmatrix}$.

4.2.2 Computation via the poles

For simplicity we take $n_1 = n_2 = n$.
Let the AR model M_1 have poles $\alpha_1, \ldots, \alpha_n$ and M_2 have poles $\beta_1, \ldots, \beta_n$. Assume the two models are stable and observable.

The cosines of the subspace angles between M_1 and M_2 are equal to the largest n eigenvalues of $\begin{pmatrix} 0 & X \\ Y & 0 \end{pmatrix} \in \mathbb{C}^{2n\times 2n}$, *where*

$$\begin{aligned} X(i,j) &= \frac{\prod_{k=1}^{n}(1-\bar{\alpha}_k\alpha_i)\prod_{k\neq i}^{n}(\beta_j-\alpha_k)}{\prod_{k\neq i}^{n}(\alpha_i-\alpha_k)\prod_{k=1}^{n}(1-\bar{\alpha}_k\beta_j)}\,, \\ Y(i,j) &= \frac{\prod_{k=1}^{n}(1-\bar{\beta}_k\beta_i)\prod_{k\neq i}^{n}(\alpha_j-\beta_k)}{\prod_{k\neq i}^{n}(\beta_i-\beta_k)\prod_{k=1}^{n}(1-\bar{\beta}_k\alpha_j)}\,, \end{aligned} \tag{6}$$

in which $A(i,j)$ denotes the element of the matrix A on the ith row and the jth column.
More details can be found in [4].

5 Relation of Martin's metric and the subspace angles between two AR models

The subspace angles between two AR models are related to the distance between AR models as defined in [14] (see also Section 3) in the following way.

Theorem 1. *For the AR models M_1 of order n_1 and M_2 of order n_2,*

$$d_M(M_1,M_2)^2 = \ln\prod_{i=1}^{n}\frac{1}{\cos^2\theta_i}\,, \tag{7}$$

where $n = \max(n_1,n_2)$ and $\theta_1, \theta_2, \ldots, \theta_n$ are the subspace angles between M_1 and M_2.

The proof can be found in [4].

6 Distance measures based on matrix norms

In this section we give two other distance measures for AR models, which are based on a definition in [8] for the distance between subspaces. We also show how they are related to the subspace angles between the models.
The distance we describe in Section 6.1 was for subspaces defined by Golub and Van Loan in [8] and is based on the matrix two-norm. We apply it to the ranges of the infinite observability matrices of the AR models. In Section 6.2 we follow the same track, but instead of the two-norm we use the Frobenius norm.

6.1 The gap between two AR models

Golub and Van Loan give in [8] a definition for the distance between two equidimensional subspaces, based on the one-to-one correspondence between subspaces and orthogonal projections. This metric is called the gap between the subspaces [7].
If P_1 is the orthogonal projection onto a subspace S_1 and P_2 the orthogonal projection onto S_2, then the gap between S_1 and S_2 is defined as

$$d_g(S_1,S_2) = \|P_1 - P_2\|_2\,. \tag{8}$$

As in the previous sections, the subspaces we focus on, are the ranges of the infinite observability matrices of the AR models.

Assume two stable and observable AR models M_1 and M_2 of order n and with infinite observability matrix $\mathcal{O}_\infty(M_1)$ and $\mathcal{O}_\infty(M_2)$ respectively, are given. For models that have a different order, the same trick is used as in Section 4.1: we add Δn poles in zero to the model with the smallest order, where Δn is the difference in model order. Let $P_1 \in \mathbb{R}^{\infty\times\infty}$ be the orthogonal projection onto the range of $\mathcal{O}_\infty(M_1)$ and $P_2 \in \mathbb{R}^{\infty\times\infty}$ the orthogonal projection onto the range of $\mathcal{O}_\infty(M_2)$.

Definition 6.1. We define the gap d_g between the AR models M_1 and M_2 as the gap between the subspaces range $(\mathcal{O}_\infty(M_1))$ and range $(\mathcal{O}_\infty(M_2))$:

$$\begin{aligned} d_g(M_1,M_2) &= d_g(\text{range}\,(\mathcal{O}_\infty(M_1))\,,\text{range}\,(\mathcal{O}_\infty(M_2))) \\ &= \|P_1 - P_2\|_2\,. \end{aligned}$$

The relation between this distance measure and the subspace angles between the AR models is

$$d_g(M_1,M_2) = \sin\theta_{\max}\,, \tag{9}$$

where $\theta_{\max}$ is the largest subspace angle between the models M_1 and M_2. The equality in (9) can also be found in [8], where it is given for general subspaces.

6.2 Distance based on the Frobenius norm

Instead of using the 2-norm of the difference between the orthogonal projections, one could as well apply the Frobenius norm:
If P_1 is the orthogonal projection onto a subspace S_1 and P_2 the orthogonal projection onto S_2, then the distance d_f between S_1 and S_2 is defined as

$$d_f(S_1,S_2) = \|P_1 - P_2\|_F\,. \tag{10}$$

We then get this new distance between AR models, which will be referred to as d_f.

Definition 6.2. For the stable and observable AR models M_1 and M_2, the distance d_f between M_1 and M_2 is defined as

$$d_f(M_1,M_2) = d_f(\text{range}\,(\mathcal{O}_\infty(M_1))\,,\text{range}\,(\mathcal{O}_\infty(M_2)))\,.$$

For this case too, there exists a relation with the subspace angles between the AR models M_1 and M_2:

$$d_f(M_1, M_2)^2 = 2\sum_{i=1}^{n} \sin^2 \theta_i \ , \tag{11}$$

where the angles $\theta_i (i = 1, \ldots, n)$ are the subspace angles between the AR models M_1 and M_2.
This can be seen as follows.
Let $U_1 \in \mathbb{R}^{\infty \times n}$ be an orthogonal matrix of which the column space is equal to range $(\mathcal{O}_\infty(M_1))$ and $U_2 \in \mathbb{R}^{\infty \times n}$ an orthogonal matrix with the same column space as $\mathcal{O}_\infty(M_2)$. The orthogonal projections onto the column spaces can then be written as $P_1 = U_1 U_1^T$ and $P_2 = U_2 U_2^T$.
For each rank n matrix M holds

$$\|M\|_F^2 = \text{trace}(M^T M) = \sum_{i=1}^{n} \sigma_i^2(M) \ ,$$

where $\sigma_i(M)$ is the ith singular value of M.
Consequently,

$$\begin{aligned} d_f(M_1, M_2)^2 &= \|P_1 - P_2\|_F^2 \\ &= \text{trace}\left((P_1 - P_2)^T (P_1 - P_2)\right) \\ &= \text{trace}(P_1) + \text{trace}(P_2) - 2\text{trace}(P_1 P_2) \\ &= 2n - 2\sum_{i=1}^{n} \sigma_i^2(U_1^T U_2) \ , \end{aligned} \tag{12}$$

where the $\sigma_i(U_1^T U_2)$ are the singular values of $U_1^T U_2$. Since the singular values of $U_1^T U_2$ are equal to the cosines of the principal angles between range $(\mathcal{O}_\infty(M_1))$ and range $(\mathcal{O}_\infty(M_2))$ (see Section 2.3), (11) follows from (12).

7 Distances and subspace angles between ARMA models

As mentioned in Section 3, Martin defined a metric, not only for AR models, but more general for ARMA models [14]. On the basis of this definition, which is repeated in Section 7.1, and a property of the distance measure, we define in Section 7.2 the subspace angles between two ARMA models. Furthermore, we will show in Section 7.3 how we can adapt Definition 6.1 and Definition 6.2 to the ARMA model class. In Section 7.4, finally, we give two methods for the computation of the subspace angles between ARMA models.

7.1 Martin's distance between ARMA models

Definition 3.1 can readily be used for the distance between ARMA models: Let M_1 and M_2 be ARMA models with cepstrum coefficients $c_1(n)$ and $c_2(n)$, $n = 0, \pm 1, \pm 2, \ldots$.

Definition 7.1. [14] The distance between M_1 and M_2 is defined as

$$d_M(M_1, M_2) = \sqrt{\sum_{n=0}^{\infty} n|c_1(n) - c_2(n)|^2} \ .$$

Since the cepstrum is the inverse Fourier transformation of the logarithm of the spectrum:

$$\ln P(z) = \ln H(z)\bar{H}(z^{-1}) = \sum_{n \in \mathbb{Z}} c_n z^{-n} \ ,$$

the following property holds [14]

$$d_M(H_1 H_3, H_2 H_3) = d_M(H_1, H_2) \ ,$$

where H_i is the transfer function of the ARMA model M_i for $i = 1, 2, 3$. This property implies that in order to compute the distance between ARMA models, it is sufficient to consider AR models: for $H_1(z) = \frac{b_1(z)}{a_1(z)}$ and $H_2(z) = \frac{b_2(z)}{a_2(z)}$, take $H_3(z) = \frac{1}{b_1(z)b_2(z)}$, so that

$$d_M\left(\frac{b_1(z)}{a_1(z)}, \frac{b_2(z)}{a_2(z)}\right) = d_M\left(\frac{1}{a_1(z)b_2(z)}, \frac{1}{a_2(z)b_1(z)}\right) . \tag{13}$$

Consider now the subspace angles between the AR models with respective transfer function $\frac{1}{a_1(z)b_2(z)}$ and $\frac{1}{a_2(z)b_1(z)}$. In accordance with Definition 4.1, they are equal to the principal angles between the range of $\begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2^{-1}) \end{pmatrix}$ and the range of $\begin{pmatrix} \mathcal{O}_\infty(M_2) & \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix}$, where the transfer function of M_1 is $\frac{b_1(z)}{a_1(z)}$, the transfer function of M_2 is $\frac{b_2(z)}{a_2(z)}$, $\mathcal{O}_\infty(M_i)$ is the infinite observability matrix of the model M_i and $\mathcal{O}_\infty(M_i^{-1})$ is the infinite observability matrix of the inverse model M_i^{-1}, for $i = 1, 2$.
We now propose the following definition of the subspace angles between ARMA models.

7.2 Subspace angles between ARMA models

Assume M_1 and M_2 are stable, minimum phase ARMA models (the poles as well as the zeros lie inside the unit circle).

Definition 7.2. We define the subspace angles between M_1 and M_2 as the principal angles between the ranges of $\begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2^{-1}) \end{pmatrix}$ and $\begin{pmatrix} \mathcal{O}_\infty(M_2) & \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix}$:

$$[M_1 \lhd M_2] = \left[\begin{pmatrix} \mathcal{O}_\infty(M_1) \end{pmatrix} \middle| \begin{pmatrix} \mathcal{O}_\infty(M_2^{-1}) \end{pmatrix} \lhd \begin{pmatrix} \mathcal{O}_\infty(M_2) \end{pmatrix} \middle| \begin{pmatrix} \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix}\right] .$$

Or,

$$\left[\frac{b_1(z)}{a_1(z)} \lhd \frac{b_2(z)}{a_2(z)}\right] = \left[\frac{1}{a_1(z)b_2(z)} \lhd \frac{1}{a_2(z)b_1(z)}\right] , \tag{14}$$

which reflects the property for the distance between two ARMA models in Equation (13).
From Definition 7.2 and Equation (13) it is clear that Theorem 1, which was given for AR models, is also valid for ARMA models:

$$d_M(M_1, M_2)^2 = \ln \prod_{i=1}^{2n} \frac{1}{\cos^2 \theta_i} \ ,$$

where M_1 and M_2 are ARMA models of order n and the angles $\theta_i (i = 1, \ldots, 2n)$ are the subspace angles between M_1 and M_2.

7.3 The gap and Frobenius norm based distances between ARMA models

The gap and Frobenius norm based distances can now be defined for ARMA models as follows:

$$d_g(M_1, M_2) = d_g\Big(\text{range}\begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2^{-1}) \end{pmatrix}, \text{range}\begin{pmatrix} \mathcal{O}_\infty(M_2) & \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix}\Big)$$

and

$$d_f(M_1, M_2) = d_f\Big(\text{range}\begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2^{-1}) \end{pmatrix}, \text{range}\begin{pmatrix} \mathcal{O}_\infty(M_2) & \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix}\Big),$$

where the definitions for these metrics for subspaces were given in (8) and (10) respectively.
The relations between the subspace angles $\theta_1, \ldots, \theta_{2n}$ between the ARMA models M_1 and M_2 and the defined distance measures also remain as follows

$$d_g(M_1, M_2) = \sin\theta_{\max} ,$$

and

$$d_f(M_1, M_2)^2 = 2\sum_{i=1}^{2n} \sin^2\theta_i .$$

7.4 Computation of the subspace angles between two stable and observable ARMA models

In this section we show how the subspace angles between two ARMA models can be computed. The two methods are completely analogous to the methods for AR models discussed in Section 4.2. In Section 7.4.1 the subspace angles are computed via the observability Gramian of a concatenation of models and inverse models. In Section 7.4.2 a method is described to compute the subspace angles via the poles and zeros of the two ARMA models.

7.4.1 Computation via the observability Gramian

Assume the nth order stable and observable SISO ARMA models M_1 and M_2 have the following state space representation:

$$\begin{cases} x_{k+1} &= A_i x_k + w_k , \\ y_k &= C_i x_k + v_k , \end{cases} \tag{15}$$

where $i = 1, 2$.
The values $y_k \in \mathbb{R}$ are the measurements at time instant k of the output of the process. The vector $x_k \in \mathbb{R}^n$ is the state vector of the process at discrete time instant k, where n is the order of the model, i.e. the number of poles. The sequences $v_k \in \mathbb{R}$ and $w_k \in \mathbb{R}^n$ are zero mean, stationary, Gaussian white noise signals with

$$E\left\{\begin{pmatrix} w_k \\ v_k \end{pmatrix}\begin{pmatrix} w_l^T & v_l^T \end{pmatrix}\right\} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}\delta_{kl} ,$$

where δ_{kl} is the Kronecker delta. The matrices $A_i \in \mathbb{R}^{n\times n}$ and $C_i \in \mathbb{R}^{1\times n}$ are the system matrices, respectively output matrices of the models. The eigenvalues of A_1 are the poles $\alpha_1, \ldots, \alpha_n$ of M_1 and the eigenvalues of A_2 are the poles $\beta_1, \ldots, \beta_n$ of M_2.
The model (15) can be converted into a so-called forward innovation form:

$$\begin{cases} \hat{x}_{k+1} &= A_i\hat{x}_k + K_i e_k \\ y_k &= C_i\hat{x}_k + e_k , \end{cases} \tag{16}$$

where $K_i \in \mathbb{R}^n$ is the Kalman gain (see [15], p. 61).
It can be shown that the forward innovation model is minimum phase, implying that the eigenvalues of $A_i - K_iC_i$, which are the zeros of the model, lie within the unit circle. We adopt the notation $M_i = [A_i, K_i, C_i]$ for the realization in (16). The state space description of the inverse model M^{-1} is then equal to

$$[A_i, K_i, C_i]^{-1} = [A_i - K_iC_i, K_i, -C_i] . \tag{17}$$

From Definition 7.2 and Section 4.2.1 follows that the subspace angles between M_1 and M_2 can be obtained from the eigenvalues of $\begin{pmatrix} 0 & Q_{11}^{-1}Q_{12} \\ Q_{22}^{-1}Q_{21} & 0 \end{pmatrix}$, where

$$\begin{aligned} Q &= \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \\ &= \begin{pmatrix} \mathcal{O}_\infty(M_1)^T \\ \mathcal{O}_\infty(M_2^{-1})^T \\ \mathcal{O}_\infty(M_2)^T \\ \mathcal{O}_\infty(M_1^{-1})^T \end{pmatrix} \times \\ &\begin{pmatrix} \mathcal{O}_\infty(M_1) & \mathcal{O}_\infty(M_2^{-1}) & \mathcal{O}_\infty(M_2) & \mathcal{O}_\infty(M_1^{-1}) \end{pmatrix} . \end{aligned}$$

The matrix Q can be regarded as the observability Gramian of the model with state space matrices

$$\begin{aligned} \mathcal{A} &= \left(\begin{array}{cc|cc} A_1 & 0 & 0 & 0 \\ 0 & (A_2 - K_2C_2) & 0 & 0 \\ \hline 0 & 0 & A_2 & 0 \\ 0 & 0 & 0 & (A_1 - K_1C_1) \end{array}\right) , \\ \mathcal{C} &= \begin{pmatrix} C_1 & -C_2 & C_2 & -C_1 \end{pmatrix} \end{aligned}$$

(see Equation (17). Consequently, the matrix Q can be obtained by solving the Lyapunov equation:

$$\mathcal{A}^T Q \mathcal{A} - Q = -\mathcal{C}^T\mathcal{C} . \tag{18}$$

The computation of the subspace angles between the ARMA models M_1 and M_2 boils down to the following steps:

1. Compute the solution $Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$ of the Lyapunov equation (18).
2. Compute the eigenvalues $\lambda_1, \ldots, \lambda_{4n}$ of $\begin{pmatrix} 0 & Q_{11}^{-1}Q_{12} \\ Q_{22}^{-1}Q_{21} & 0 \end{pmatrix}$,
3. The $2n$ largest eigenvalues λ_i are equal to the cosines of the subspace angles $\theta_1, \ldots, \theta_{2n}$ between M_1 and M_2.

7.4.2 Computation via the poles and zeros of the ARMA models

The subspace angles between two ARMA models M_1 with transfer function $\frac{b_1(z)}{a_1(z)}$ and M_2 with transfer function $\frac{b_2(z)}{a_2(z)}$ are equal to the subspace angles between the AR models M_3 with transfer function $\frac{1}{a_1(z)b_2(z)}$ and M_4 with transfer function $\frac{1}{a_2(z)b_1(z)}$ (see Equation (14)). Consequently, the poles of the AR models M_3 and M_4 are known if we know the poles and zeros of the ARMA models M_1 and M_2 and then we can apply the method described in Section 4.2.2.

8 Conclusions

In this paper we have proposed a definition for the subspace angles and directions between two ARMA models and we have shown a relation between these angles and a recently defined distance measure for ARMA models [14].
In the near future, these new notions of distance and angles between models will be applied to several engineering applications, such as signal classification, fault detection, the calculation of so-called stabilization diagrams in vibrational analysis etc ...
Many questions remain to be tackled. Future developments will comprise the extension to MIMO (multiple input multiple output) models and to deterministic systems. Furthermore, the apparent relation with the notion of mutual information will be explored.

Acknowledgments

Bart De Moor is a Research Associate with the F.W.O. (Fund for Scientific Research-Flanders) and professor extraordinary at the K.U.Leuven. Katrien De Cock is a Research Assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry). This work is supported by several institutions: the Flemish Government (Research Council K.U.Leuven: Concerted Research Action Mefisto-666, the FWO projects G.0240.99, G.0256.97, and Research Communities: ICCoS and ANMMM, IWT projects: EUREKA 1562-SINOPSYS, EUREKA 2063-IMPACT, STWW), the Belgian State, Prime Minister's Office (IUAP P4-02 (1997-2001) and IUAP P4-24 (1997-2001), Sustainable Mobility Program - Project MD/01/24 (1997-2000)), the European Commission (TMR Networks: ALAPEDES and System Identification, Brite/Euram Thematic Network: NICONET), Industrial Contract Research (ISMC, Data4S, Electrabel, Laborelec, Verhaert, Europay)

References

[1] Akaike H. *Stochastic theory of minimal realization.* IEEE Transactions on Automatic Control, Vol.19, pp.667–674, 1974.

[2] Björck Å. and Golub G. H. *Numerical methods for computing angles between linear subspaces.* Mathematics of computation, Vol.27, pp.579–594, 1973.

[3] Caines P. E. *Linear stochastic systems*. Wiley New York, 1988.

[4] De Cock K. and De Moor B. *Subspace angles between linear stochastic models*[1]. Technical Report TR 99–101a, ESAT-SISTA,K.U.Leuven (Leuven, Belgium), 2000. Submitted for publication.

[5] Gelfand I. M. and Yaglom A. M. *Calculation of the amount of information about a random function contained in another such function.* American Mathematical Society Translations, series 2, Vol.12, pp.199–246, 1959.

[6] Gittins R. *Canonical Analysis; A review with applications in ecology*, Vol.12 of *Biomathematics*. Springer-Verlag, Berlin, 1985.

[7] Gohberg I., Lancaster P. and Rodman L. *Invariant subspaces of matrices with applications*. Wiley New York, 1986.

[8] Golub G.H. and Van Loan C.F. *Matrix computations*. The Johns Hopkins University Press, second edition, 1989.

[9] Golub G.H. and Zha H. *The canonical correlations of matrix pairs and their numerical computation.* In Adam Bojanczyk and George Cybenko, editors, *Linear algebra for signal processing*, Vol.69, pp.59–82. Springer-Verlag, New York, 1995.

[10] Hotelling H. *Relations between two sets of variates.* Biometrika, Vol.28, pp.321–372, 1936.

[11] Jordan C. *Essai sur la géométrie à n dimensions.* Bulletin de la Société Mathématique, Vol.3, pp.103–174, 1875.

[12] Kailath T. *A view of three decades of linear filtering theory.* IEEE Transactions on Information Theory, Vol.20, No.2, pp.146–181, 1974.

[13] Martin R.J. *Autoregression and cepstrum-domain filtering.* Signal Processing, Vol.76, pp.93–97, 1999.

[14] Martin R.J. *A metric for ARMA processes.* IEEE Transactions on Signal Processing, Vol.48, No.4, pp.1164–1170, 2000.

[15] Van Overschee P. and De Moor B. *Subspace identification for linear systems: Theory – Implementation – Applications*. Kluwer Academic Publishers, 1996.

[1] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* as file *pub/sista/decock/reports/99–101.ps.gz*