AN ALGORITHM FOR CONSTRAINED ORTHOGONALIZATION

Ivanil S. Bonatti, Pedro L. D. Peres and Amauri Lopes

School of Electrical and Computer Engineering
University of Campinas
CP 6101, 13081-970, Campinas - SP - Brazil.
e-mails: {*ivanil@dt,peres@dt,amauri@decom*}*@fee.unicamp.br*

ABSTRACT

This paper presents a procedure for the generation of an orthogonal set of functions, starting from a set of linearly independent non-orthogonal functions. The main characteristic of the algorithm is to produce the minimal perturbation on the original set.

Keywords: Orthonormal functions; Gram-Schmidt orthonormalization; Signal representation; Orthonormal bases.

1. INTRODUCTION

Bases for signal representation are perhaps the most important tool for signal processing and analysis [1], [2], [3], [5], [6], [7], [8], [9], [10]. The action of decomposing a signal through a base formed by a set of suitable signals is present in almost all procedures currently adopted to convey or to extract information. The orthogonal bases are massively employed mainly because they allow a simple decoupled computation of the coefficients of a signal representation.

The well known Gram-Schmidt procedure generates orthogonal functions starting from a set of linearly independent functions. However, the procedure leads to new functions whose shapes can be very different from the original ones.

A new procedure to orthogonalize a set of linearly independent functions is proposed in this paper. Its main characteristic is to preserve, subject to some criterion, the shapes of the original base components. A comparison with the classical Gram-Schmidt procedure is provided.

2. GRAM-SCHMIDT

Consider the set $\mathcal{N}$ of linearly independent real functions $f_n(t)$, $n = 1, \ldots, N$ with finite energy. A signal $y(t)$ not necessarily belonging to the space $\mathcal{S}_N$ spanned by the set $\mathcal{N}$ can be approximately expressed as

$$y(t) \cong \sum_{n=1}^{N} \alpha_n f_n(t) \qquad (1)$$

The coefficients α_n that minimize the mean square error

$$\left\langle \epsilon^2(t) \right\rangle = \int_{-\infty}^{+\infty} \left[y(t) - \sum_{n=1}^{N} \alpha_n f_n(t) \right]^2 dt \qquad (2)$$

are given by

$$\alpha = R^{-1} \left\langle f y(t) \right\rangle \qquad (3)$$

where $\alpha \in \mathbb{R}^N$ and $f = [f_1(t) \ldots f_N(t)]'$ are column vectors. The correlation matrix $\left\langle R = f f' \right\rangle$ is composed by the elements

$$r_{kl} = \left\langle f_k(t) f_l(t) \right\rangle = \int_{-\infty}^{+\infty} f_k(t) f_l(t) dt \text{ for } k, l = 1, \ldots, N \qquad (4)$$

Note that the calculation of each coefficient α_n simultaneously involves all the $f_n(t)$ functions. If the set $\mathcal{N}$ is composed by orthogonal functions then R is a diagonal matrix and each coefficient α_n can be calculated only from the corresponding $f_n(t)$ function and the signal $y(t)$.

To orthogonalize the set $\mathcal{N}$, consider the non-singular matrix $Q \in \mathbb{R}^{N \times N}$ which produces the linear transformation

$$g \triangleq Q f \qquad (5)$$

Note that the vector of functions g produces a new base that spans the same space $\mathcal{S}_N$ that the one generated by f. The orthonormalization requires that $\left\langle g g' \right\rangle = \mathbf{I}$ yielding

$$Q R Q' = \mathbf{I} \qquad (6)$$

Equation (6) can be viewed as a system of quadratic equations with N^2 unknown variables and $N(N+1)/2$ constraints, since by construction R is a symmetric positive definite matrix [5]. As a consequence, there are several different ways to generate the orthonormal base g from the given set of linearly independent functions f.

The symmetry and the positive definiteness characteristics of R allow to obtain a solution of equation (6) by applying the Cholesky factorization to R, producing a lower triangular matrix L such that $R = LL'$ [4], yielding

$$QRQ' = (QL)(QL)' = \mathbf{I} \tag{7}$$

The Cholesky factorization induces $Q = L^{-1}$ as a solution of (6), which is the classical Gram-Schmidt orthonormalization.

To illustrate, consider the set of five linearly independent triangular functions $f_n(t) = f(t-nT)$ for $n = 1, \dots, 5$ where

$$f(t) = \begin{cases} (t/T) + 1 \text{ for } -T \leq t < 0 \\ -(t/T) + 1 \text{ for } 0 \leq t \leq T \\ 0 \quad \text{elsewhere} \end{cases} \tag{8}$$

with $T = 1.5$ as illustrated in the figure 1.

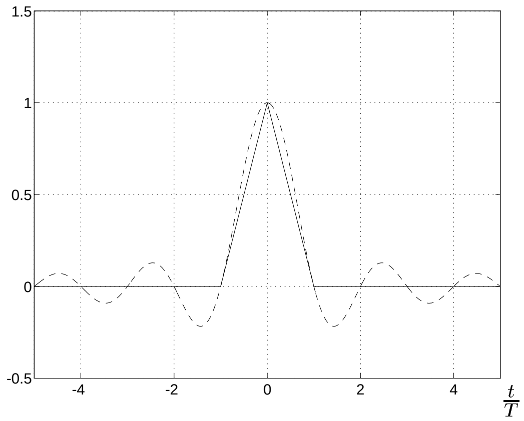

Figure 1: Triangular pulse that approximates the sampling function Sa$(\pi t/T)$.

The correlation matrix R (which is a tri-diagonal matrix) and the corresponding linear transformation $Q = L^{-1}$ for this example are given by

$$R = \begin{bmatrix} 1 & 0.25 & 0 & 0 & 0 \\ 0.25 & 1 & 0.25 & 0 & 0 \\ 0 & 0.25 & 1 & 0.25 & 0 \\ 0 & 0 & 0.25 & 1 & 0.25 \\ 0 & 0 & 0 & 0.25 & 1 \end{bmatrix} \tag{9}$$

$$Q = \begin{bmatrix} +1.000 & 0 & 0 & 0 & 0 \\ -0.258 & +1.033 & 0 & 0 & 0 \\ +0.069 & -0.276 & +1.035 & 0 & 0 \\ -0.019 & +0.074 & -0.277 & +1.035 & 0 \\ +0.005 & -0.019 & +0.074 & -0.277 & +1.035 \end{bmatrix} \tag{10}$$

Figure 2 shows the orthonormalized functions g. As it is well known, the classical Gram-Schmidt orthonormalization procedure is such that $g_1(t) = f_1(t)$ but $g_n(t) \neq f_n(t)$ for $n = 2, \dots, N$ and the changes on the original shapes are more important as n increases.

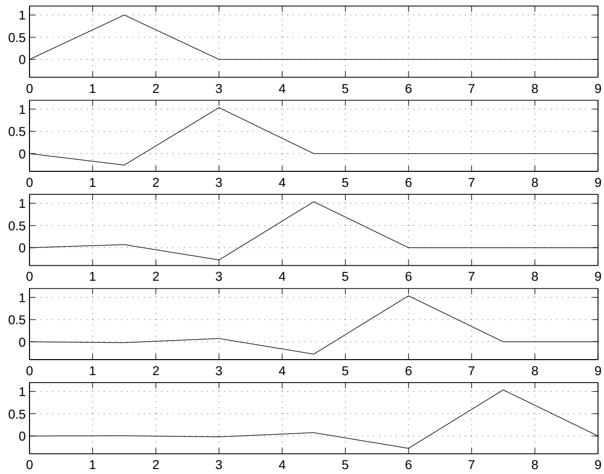

Figure 2: Functions g produced by the Gram-Schmidt orthonormalization procedure applied on the triangular functions f.

3. CONSTRAINED ORTHOGONALIZATION

The aim of the procedure proposed here is, of course, to orthogonalize the original set but obtaining a set of functions that best preserves the shapes of the original functions. For that, a quadratic criterion is used.

The problem becomes

$$\min_{g = Qf;\, g \text{ orthogonal}} \sum_{n=1}^{N} \int_{-\infty}^{+\infty} [g_n(t) - f_n(t)]^2 \, dt \tag{11}$$

Let the vectors f_Δ and g_Δ be the functions sampled with a period Δ, yielding the approximative numerical computation for the correlation matrix R:

$$R \cong f_\Delta f'_\Delta \Delta \tag{12}$$

Therefore, problem (11) can be written as

$$\min_{g_\Delta = Qf_\Delta;\, \Delta g_\Delta g'_\Delta = \mathbf{I}} \text{Tr}\left[(g_\Delta - f_\Delta)(g_\Delta - f_\Delta)'\right]\Delta \tag{13}$$

$$\min_{QRQ' = \mathbf{I}} \text{Tr}\left[QRQ' + R - QR - RQ'\right] \tag{14}$$

where $\text{Tr}(\cdot)$ stands for trace of the square matrix argument, that is, $\text{Tr}(M) = \sum_{i=1}^{n} m_{ii}$ for $M \in \mathbb{R}^{n\times n}$.

The solution to problem (14) also solves

$$\max_{QRQ' = \mathbf{I}} \ \text{Tr}\,[2RQ] \tag{15}$$

which involves the maximization of a linear cost function subject to quadratic constraints. Since R is a positive definite matrix, this problem has only one solution, which can be obtained using the Lagrange multiplier Y.

The Lagrangian function is given by

$$l(Q,Y) = Tr[2RQ + Y'(QRQ' - \mathbf{I})] \tag{16}$$

yielding the stationary conditions [11]

$$QRQ' = \mathbf{I} \quad ; \quad 2R + YQR + Y'QR = 0 \tag{17}$$

Note that the constraint $QRQ' = \mathbf{I}$ is symmetric implying that $Y' = Y$ and $Y = -Q^{-1}$ which, applied to $QRQ' = \mathbf{I}$ leads to $Y^2 = R$. Thus, the optimal solution of (15) is $Q = R^{-0.5}$.

As R is a real symmetric positive definite matrix

$$Q = R^{-0.5} = U\Lambda^{-0.5}U' \tag{18}$$

where U is a unitary matrix (i.e. $U'U = \mathbf{I}$) and Λ is a diagonal matrix formed by the eigenvalues of R [4].

Observe that the solution Q is a symmetric matrix. Indeed, by imposing this condition on equation (6) and applying the Schur decomposition on R [4], one gets

$$R = U\Lambda U' \implies Q = U\Lambda^{-0.5}U' \triangleq R^{-0.5} \tag{19}$$

Figure 3 shows the proposed constrained orthonormalization applied to the triangular functions given in (8). Matrix Q is given by

$$Q = \begin{bmatrix} +1.026 & -0.136 & +0.027 & -0.006 & +0.001 \\ -0.136 & +1.053 & -0.142 & +0.029 & -0.006 \\ +0.027 & -0.142 & +1.054 & -0.142 & +0.027 \\ -0.006 & +0.029 & -0.142 & +1.053 & -0.136 \\ +0.001 & -0.006 & +0.027 & -0.136 & +1.026 \end{bmatrix} \tag{20}$$

Note that the symmetry of matrix Q results in a small perturbation equally distributed in all the original functions.

4. A TWO-DIMENSIONAL EXAMPLE

A better geometric understanding of the method proposed is illustrated through a simple example in a Cartesian plan.

Consider the matrix $F \in \mathbb{F}^{2\times 2}$ composed by two linearly independent row vectors and the corresponding correlation matrix R given by

$$F = \begin{bmatrix} 0 & 1 \\ 0.707 & 0.707 \end{bmatrix} \quad ; \quad R = \begin{bmatrix} 1 & 0.707 \\ 0.707 & 1 \end{bmatrix} \tag{21}$$

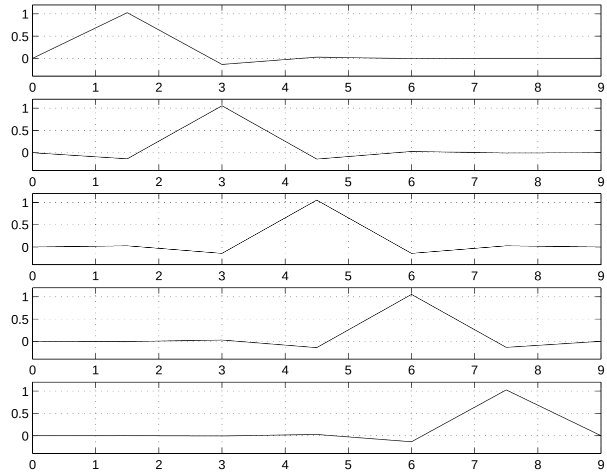

Figure 3: Functions g produced by the constrained orthonormalization procedure applied on the triangular functions f.

The linear transformations that produce the Gram-Schmidt and the constrained orthonormalizations are given respectively by

$$\begin{bmatrix} 1 & 0 \\ -1 & 1.414 \end{bmatrix} \quad ; \quad \begin{bmatrix} 1.307 & -0.541 \\ -0.541 & 1.307 \end{bmatrix} \tag{22}$$

and the two resulting orthonormal sets g (i.e. unitary matrices) are respectively

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad ; \quad \begin{bmatrix} -0.383 & 0.924 \\ 0.924 & 0.383 \end{bmatrix} \tag{23}$$

Figure 4 shows that the resulting vectors in the constrained orthogonalization are equally perturbed with respect to the original vectors whereas in the Gram-Schmidt procedure the first vector is preserved and the second one is modified.

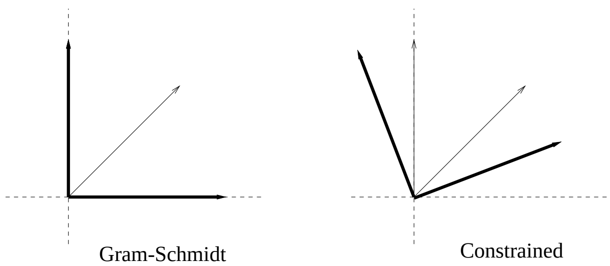

Figure 4: Resulting vectors for the two dimensional example.

5. CONCLUSION

A new procedure to orthogonalize a set of linearly independent functions has been presented. The algorithm is similar to the classical Gram-Schmidt one, but presents the additional feature of minimizing the total mean square difference between the original functions and the corresponding orthogonal ones.

Acknowledgement

This research has been supported by grants from "Conselho Nacional de Desenvolvimento Científico e Tecnológico" — CNPq, Brazil.

6. REFERENCES

[1] C. S. Burrus, R. A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelet Transformers: A Primer*. Prentice-Hall, Englewood Cliffs, NJ, 1998.

[2] C. T. Chen. *System and Signal Analysis*. Saunders College Publishing, Orlando, FL, 1989.

[3] C. K. Chui. *Wavelets: A Mathematical Tool for Signal Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997.

[4] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins Press, New York, 1996.

[5] S. Haykin. *Digital Communications*. John Wiley & Sons, New York, NY, 1988.

[6] B. P. Lathi. *Signals, Systems and Communication*. John Wiley & Sons, New York, NY, 1965.

[7] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

[8] A. V. Oppenheim, A. S. Willsky, and I. T. Young. *Signals and Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1983.

[9] L. R. Rabiner and R. W. Schafer. *Digital Processing of Speech Signals*. Prentice-Hall, Englewood Cliffs, NJ, 1978.

[10] C. W. Therrien. *Discrete Random Signals and Statistical Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1992.

[11] K. Zhou, K. Glover, and J. C. Doyle. *Robust and Optimal Control*. Prentice Hall, USA, 1995.