QR DECOMPOSITION FOR ADAPTIVE FILTERING APPLICATION

FRIDAY ZINZENDOFF OKWONU

SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITI SAINS MALAYSIA

MAY 2011

QR DECOMPOSITION FOR ADAPTIVE FILTERING APPLICATION

by

FRIDAY ZINZENDOFF OKWONU

Thesis submitted in fulfillment of the requirements
for the degree of Master of Science

May 2011

Acknowledgement

This thesis is dedicated to Almighty GOD for his kindness and love. My profound gratitude goes to my supervisor Dr. Noor Atinah Ahmad for her diligent, unreserved support and consistent advice. Special thanks to the dean, Deputy Dean and staff of School of Mathematical Sciences. My profound gratitude goes to my father Chief J. O. Okwonu and brothers especially Engr. and Mrs. B. O. Okwonu for his love and care. My gratitude goes to Christopher Ajiri, Favour Okwonu, Engr. M. E. Okwonu, G. O. Okwonu, Ochuko Ejiagu, Efe Emejor and Johnson Grace. I deeply appreciate my friends and mentor Obarisi of Orogu Kingdom Chief Ovie Omo-Agege, Friday Onodje, Prof. Okpako J. Enawho (JP), Prof. P. N. Okonta and family, Sir and Lady J. S. Apanapudor, Dr. F. O. Arunayi and family, Prof. C. O. Orobu and family, Mr. and Mrs. Legborsi K. Obenade, Major Chukuma Kaduna Nzeogwu, Christopher Omobo, Friday Ikikiru, Sunny S. Shola, Mahshid Mosherefi.

Table of Contents

List of Tables

List of Figures

PENGURAIAN QR UNTUK PENAPIS MUDAH-SUAI

Abstrak

Tesis ini bertujuan mengkaji masalah penapisan mudah suai (*adaptive filtering*) berdasarkan teknik penguraian QR. Penapis mudah suai ialah suatu penapis digital yang boleh melaras pekalinya untuk meminimumkan fungsi ralat yang ditakrifkan. Algoritma mudah suai digunakan bagi menyesuaikan pekali daripada penapis yang digunakan dalam proses tidak tetap, yang pekali penapis mudah suai disesuaikan untuk meminimumkan fungsi ralat. Masalah penapisan mudah suai merupakan suatu bentuk daripada masalah kuasa dua terkecil *(least squares).* Kaedah kuasa dua terkecil rekursif (*recursive least squares, RLS*) mampu mengemas kini songsangan matriks autokorelasi secara rekursif melalui lema songsangan matriks untuk mengkomput vektor pekali dan ralat yang berkaitan. Kami menggunakan penguraian QR berdasarkan putaran Givens untuk mengkaji masalah penapisan mudah suai. Putaran Givens digunakan pada algoritma penapisan mudah suai kerana sifat lelarannya yang memudahkan pengemaskinian data matriks segitiga. Kuasa dua terkecil rekursif melalui penguraian QR (*QRD–RLS)* menjelma matrik data kepada matriks segitiga atas dan mengemaskininya secara rekursif. Penjelmaan menghasilkan persamaan normal dalam bentuk yang lebih mudah dan boleh diselesaikan untuk mencari pekali vektor melalui gantian ke belakang. Sebaliknya, kuasa dua terkecil rekursif melalui penguraian QR songsang *(IQRD-RLS)* mengemas kini songsangan matriks segitiga atas, yang akhirnya membolehkan pekali vektor dikomput secara terus, iaitu, tanpa gantian ke belakang. Kami mengkaji prestasi bandingan algoritma dengan merujuk kepada syarat matriks autokorelasi masalah. Purata norma ralat pemberat digunakan untuk menganalisis kadar penumpuan dan salah-laras. Simulasi menunjukkan bahawa bagi nombor syarat yang rendah, RLS menumpu cepat dengan salah-laras setara dengan kaedah QR. Walau

bagaimanapun, apabila nombor syarat semakin bertambah, RLS menunjukkan pengurangan boleh-kesan dan menghasilkan salah-laras yang semakin tinggi. Sebaliknya teknik penguraian QR menumpu perlahan tetapi apabila nombor syarat meningkat salah-aras kekal sama. Dapatan kajian menunjukkan bahawa walaupun QRD-RLS dan IQRD-RLS menumpu pada kadar yang perlahan, namun kedua-duanya mampu mengesan isyarat-masuk pada suatu kadar yang seragam.

QR DECOMPOSITION FOR ADAPTIVE FILTERING APPLICATION

Abstract

This thesis is designed to investigate adaptive filtering problem based on QR decomposition techniques. An adaptive filter is a self modifying digital filter that adjusts its parameters in order to minimize a defined error function. Adaptive algorithm is applied to adapt the coefficient of the used filter to nonstationary process in which the coefficient of the adaptive filter is adapted in order to minimize the error function. Adaptive filtering problem is an adaptive form of least squares problem. Recursive least squares (RLS) method recursively update the inverse of the autocorrelation matrix via matrix inversion lemma in order to compute coefficient vector and associated errors recursively. We apply QR decomposition based on Givens rotations to investigate adaptive filtering problem. Givens rotations is applied to adaptive filtering algorithm because of its iterative nature that allows easy update of the triangularized data matrix. QR decomposition of recursive least squares method (QRD-RLS) transforms data matrix to upper triangular matrix and recursively update matrix. The transformation results in a reduced form of the normal equation which can be solved for the coefficient vector via backward substitution. On the other hand, inverse QR decomposition of recursive least squares method (IQRD-RLS) updates the inverse of the upper triangular matrix and desired signal vector so that the coefficient vector can be computed directly, i.e., without backward substitution. We study the comparative performance with respect to the conditioning of the autocorrelation matrix of the problem. The mean weight error norm in used to analyze the rate of convergence and misadjustment. Simulation show that for lower condition number RLS converges fast with misadjustment comparable to the QR based methods. However, as the condition number increases RLS show evidence of reduced tractability and produce high misadjustment. On the

other hand, QR decomposition based techniques converges slow but as the condition number increases misadjustment remain unchange. These results show that although QRD-RLS and IQRD-RLS converge at a slower rate, they are able to track incoming signal at a steady rate.

Chapter 1

Introduction

Filtering is a signal processing operations with diverse objectives. From mathematical point of view, filtering is a function approximation technique. Adaptive filter may be understood as self modifying digital filter that adjust its coefficients in order to minimize a predefined error function. The error function is the difference between the desired signal $d(k)$ and the adaptive filter output $y(k)$. Adaptive filters are time varying since their parameters are continually changing in order to meet performance specification. Adaptive algorithm minimizes error function which includes data matrix, desired signal and adaptive filter output signal. There are three techniques identified to derive recursive algorithms for adaptive filter operations. These can be realize via Wiener filter theory, Kalman filter theory and the method of least squares. Adaptive filtering algorithm derived based on Wiener filter theory and Kalman filter theory have their origin in a statistical formulation of the problem (Farhang - Boroujeny B. 1999). Adaptive least squares is derived from conventional least squares problem. The method of least squares is based on deterministic formulation. Adaptive algorithm is applied to adapt the coefficients of the used filter to nonstationary process; the coefficient of the filter is adapted in a process that the error signal is minimized. Adaptive filtering is applied to process and analyze electrocardiogram (ECG) and other biomedical signals (Md. Zia Ur Rahman 2009). Adaptive filtering find application in satellite communications, voice communications with control system and speech and singing voice signals (de Pavia R. C. D. 2007). It is also applied to noise cancellation and arrhythmia detection (Thakor Vv. 1991).

1.1 Problem and Methodology

We investigate the stability of QR decomposition based algorithms for adaptive filtering in comparison with the conventional recursive least squares (RLS) algorithm. To investigate this problem, the following algorithms are applied; RLS, QRD-RLS and inverse QRD-RLS. The RLS algorithm use matrix inversion lemma to update the inverse of the autocorrelation matrix and subsequently compute the associated errors. The QRD-RLS algorithm decompose the data matrix into orthogonal and upper triangular matrix, the coefficient vector is computed using backward substitution. Inverse QRD-RLS update the inverse of the upper triangular matrix using generalized matrix inversion lemma. The coefficient vector is computed directly without using backward substitution. The associated errors for the QRD techniques are computed using transformed desired signal vector. The application of QR decomposition to recursive least squares allows QR decomposition recursive least squares to be numerically stable and robust. Simulation is designed using mean weight error norm to compare the comparative performance of these algorithms in terms of convergence, stability and misadjustment.

1.1.1 Description of Methodology

In this thesis, recursive algorithm for adaptive filtering is applied to least squares method. The method of least squares may be realized via block estimation or recursive estimation. The approach based on block estimation updates the input signal on a block by block basis while recursive estimation updates the input signal on a sample by sample basis. This thesis discusses the method of least squares with particular interest on recursive estimation. Recursive estimation approach includes recursive least squares and QR decomposition techniques.

Many problems in signal processing can be formulated as least squares problem

$$\min_{\mathbf{w}} \| \mathbf{Xw} - \mathbf{d} \|^2, \qquad (1.1.1)$$

the data matrix $\mathbf{X}$ is an $k \times n$, $\mathbf{d}$ is an $k \times 1$ vector and $\mathbf{w}$ is $n \times 1$ unknown vector to be estimated and ||.|| is the Euclidean norm. Least squares solution minimizes the sum of squared residual, if the rank of the data matrix is less than n then the solution to least squares is not unique (Bjorck A. 1996). The basic computational tool to solve least squares problem introduced by Gauss was to form and solve normal equation.

$$\mathbf{X}^T \mathbf{Xw} = \mathbf{X}^T \mathbf{d}, \qquad (1.1.2)$$

this was solved by symmetric Gaussian elimination technique. Research on how to obtain reliable and simple way of solving normal equation continues not until after 1924 when Cholesky factorization was introduced. Least squares problem can be solved by normal equation. Forming normal equation and solving it is not recommended in general (Na Li 2006) because it suffers various degree of numerical difficulties. Forming normal equation may require squaring the condition number of the original problem (Golub G. H. 1989). Least squares method approaches the problem of filter optimization from deterministic point of view in which the cost function is the sum of weighted error squared for the given data. Least squares method is used to approximately solve an over determine system of equation in which there are more equations than unknown.

Adaptive filtering problem is an adaptive form of least squares problem. The parameters of adaptive least squares changes as new input samples are received. The cost function of adaptive least squares unlike conventional least squares is adaptable to new input samples. The formulation of least squares problem as adaptive least squares problem is to ensure adaptability of the least squares solution when new input vectors are received such that it corresponds to adaptive filtering process. Adaptive filtering is one special aspect that adaptive least squares problem can be applied. The techniques used to compute adaptive least squares problem is similar to those applied to compute conventional least squares problem. Adaptive filter employ recursive least squares technique to minimize error function. In this thesis, recursive formulation of least squares problem that updates adaptive filter tap weight coefficient vector after the arrival of every sample input vector is considered.

Methods for computing adaptive least squares problem may be classified as direct or iterative. One special technique is the direct method based on QR decomposition (QRD) which adapt the recursive structure of QR decomposition to produce orthogonal factorization of adaptive least squares problem.

Direct methods are very robust and requires predictable amount of resources in terms of time and storage (Benzi M. 2002). Unfortunately, direct methods scale poorly with large problem size and memory requirements. Direct techniques include the following: Gaussian elimination, Cholesky factorization, LU decomposition and QR decomposition; of these techniques we apply QR decomposition.

Iterative methods require little storage and often require few operations than direct methods. However iterative methods don't have the reliability enjoyed by direct methods. In some applications, iterative methods often fail and preconditioning is necessary and do not suggest to attain convergence within a stipulated time. Iterative methods includes but to mention a few; Steepest Descent (SD), Conjugate Gradient (CG) and Least Mean Square (LMS).

Recursive least squares (RLS) is a modified version of conventional least squares problem. When solutions to least squares problem are computed and updated each time new input samples arrive the solution to the system becomes recursive. RLS updates the estimate of least squares minimization problem. The computational procedure of RLS begins with unknown data value or initial condition and applies the new data sample to update the previous data value. RLS is often described as time varying process since its parameters are recursively updated when new sample arrives. RLS recursively update solution to linear least squares filter in which the inverse of the autocorrelation matrix is recursively updated via matrix inversion lemma. The recursiveness of recursive least squares corresponds to adaptive filtering application. RLS solve adaptive filtering problem in order to compute coefficient vector and associated errors recursively. RLS depend heavily on input signal vector. RLS has excellent performance when working in time varying environment than stationary environment. The acceptance of the RLS algorithm has been impeded by unacceptable numerical performance in limited precision environment (Alexander S.T. 1993). The parameters used in deriving recursive least squares (RLS) are autocorrelation matrix $\mathbf{U}(k)$ and cross correlation vector $\mathbf{p}(k)$. These parameters are obtained based on normal equation for linear least squares filter $(\mathbf{U}(k)\mathbf{w}(k) = \mathbf{p}(k))$. To compute and update the coefficient vector recursively we apply matrix inversion lemma to recursively update the inverse of the autocorrelation matrix $\mathbf{U}^{-1}(k)$ based on solution to linear least squares filter. RLS is a special form of Kalman filter which is a particular type of least squares estimation. Recursive approach to compute solution to least squares problem was proposed by Gentleman and Kung (Regalia P. 2009). RLS algorithm was introduced in 1950 by

Placket (1950), following Placket (1950) work closely Godard (1974) used the application of Kalman filter theory to obtain RLS. However, Sayed and Kailath (1994) further expatiated on the relationship between RLS and Kalman filter theory for solving linear adaptive filtering problem. For RLS algorithm to operate in time varying environment the forgetting factor λ should be less than one, this allows the RLS algorithm to utilize finite memory. In this regard, RLS has the capability to track signal variation slowly. When the forgetting factor is less than unity, the adaptive filter coefficient is inconsistent (Haykin S. 1991). This process causes noise in the coefficient of the adaptive filter, with result that they become misadjusted from their optimum setting. The above discussion occurs if the forgetting factor is less than unity. The conditioning of the autocorrelation matrix of the problem determines the rate of convergence of the algorithm (see Figure 7.2). Recursive least squares (RLS) converges fast for low condition number. However as the condition number increases RLS show evidence of reduced tractability and produce high misadjustment, see Figure 7.2 for detail. Simulation shows that RLS is numerically unstable (Figure 7.2). In order to resolve the numerical instability associated with RLS an orthogonal technique (QR decomposition) is applied directly to transform the data matrix to orthogonal matrix $\mathbf{Q}$ and upper triangular matrix $\mathbf{R}$.

QR algorithm was proposed by Francis J. G. F. (1961) to modify LR algorithm (LR algorithm transform data matrix to lower and upper triangular matrix) proposed by Rutisauser and Schwarz (1958) (Francis J. G. F. 1961). Due to numerical instability associated with recursive least squares, QR decomposition was proposed (Alexander S.T. 1993). In adaptive filtering, QR decomposition apply time recursive in order to accept input data and desired signal at time instant k. Conventional QR decomposition transform data matrix to orthogonal and upper triangular matrix and also transform desired signal vector. Conventional QR decomposition can be described as follow

1. Transform data matrix $\mathbf{X}$ into orthogonal matrix $\mathbf{Q}$ and upper triangular matrix $\mathbf{R}$, the data matrix is an $k \times n, k \geq n$ is a full rank. Where $\mathbf{Q}$ is $k \times k$, $\mathbf{R}$ is $k \times n$ upper triangular matrix,

$$\mathbf{Q}^T \mathbf{X} = \mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix},$$

   the matrix $\mathbf{R}_1$ is an $n \times n$ upper triangular matrix (full rank since the data matrix is full rank), $\mathbf{0}$ is $(k-n) \times n$ null matrix.

2. Transform desired signal vector $\mathbf{d}$

$$\mathbf{Q}^T \mathbf{d} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix},$$

   the vector $\mathbf{c}_1$ is $n \times 1$ and $\mathbf{c}_2$ is $(k-n) \times 1$ vector.

3. Apply backward substitution to compute coefficient vector.

Recursive QR decomposition (QRD) decomposes data matrix into orthogonal matrix $\mathbf{Q}(k)$ and upper triangular matrix $\mathbf{R}(k)$ with respect to time, where k is the time index. Recursive QR decomposition can be described as follow

1. Suppose that $\mathbf{Q}(k)$ and $\mathbf{R}(k)$ are the QR factor of the data matrix $\mathbf{X}(k)$, the data matrix is $k \times n$, $k \geq n$, $\mathbf{Q}(k)$ is $k \times k$ orthogonal matrix, $\mathbf{R}(k)$ is $k \times n$ upper triangular matrix such that

$$\mathbf{Q}^T(k)\mathbf{X}(k) = \mathbf{R}(k) = \begin{bmatrix} \mathbf{R}_1(k) \\ \mathbf{0} \end{bmatrix}.$$

2. If additional data is received the data matrix becomes $\mathbf{X}(k+1)$ with dimension $(k+1) \times n$, this approach is accompanied by introducing additional row $\mathbf{x}^T(k+1)$ such that the QR factor of the data matrix $\mathbf{X}(k+1)$ can be expressed as

$$\mathbf{X}(k+1) = \begin{bmatrix} \mathbf{X}(k) \\ \mathbf{x}^T(k+1) \end{bmatrix},$$

$$\mathbf{Q}^T(k)\mathbf{X}(k+1) = \begin{bmatrix} \mathbf{R}(k) \\ \mathbf{x}^T(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1(k) \\ \mathbf{0} \\ \mathbf{x}^T(k+1) \end{bmatrix}.$$

3. Transform the desired signal vector recursively.

4. Apply backward substitution to recursively compute coefficient vector.

Orthogonal techniques include the following; Givens rotations, Householder transformation and Gram Schmdit orthogonalization. In this thesis, we apply QR decomposition (QRD) based on Givens rotations. Givens rotations technique is used to rotate individual vector via fixed angle which is denoted as matrix with parameters $[c,s]$ (Lodha N. 2009). Givens rotations zeros out element step by step (one at a time) unlike Householder transformation that zero multiple element per reflection (Diniz P. S. R. 2002; Apolinario J. A. 2009; Dimpesh P. 2009). Givens rotations is a vital technique in adaptive filtering algorithm because its iterative in nature as it corresponds to recursive least squares (RLS) in order to update and triangularize the data matrix recursively and perform rank one update. The blend of these recursiveness makes this system different from conventional least squares problem as such solutions obtained are recursive in order to correspond to adaptive filtering application. QR decomposition (QRD) using Givens rotations can easily be mapped into systolic array structure for parallel implementation (Tjoa S. 2005; Diniz P. S. R. 2008). Gentleman proposed QR decomposition to solve RLS (de Campos M. L. R. 2009; Gentleman W. M. 1981). He applied triangular array to avoid matrix inversion and employ pipeline sequence via Givens rotations to perform backward substitution (Apolinario (Jr.) J. A. 2009). QR decomposition recursive least squares was proposed by McWhiter to analyze systolic via Givens rotations by performing QR decomposition on the input data matrix to compute residual error without backward substitution (Apolinario (Jr.) J. A. 2009). The idea about QRD-RLS via Givens rotations is that after the data matrix has been transformed to upper triangular matrix we apply sequence of Givens rotations to update the upper triangular matrix and the desired signal vector. Backward substitution is used to compute the coefficient vector. The associated errors (a priori and a posteriori errors) can be computed based on Givens cosine term and transformed desired signal vector. The QR decomposition for recursive least squares is numerically stable and produce stable misadjustment (Figure 7.3).

Inverse QR decomposition recursive least squares (IQRD-RLS) algorithm was proposed to compute the coefficient vector directly (Alexander S.T. 1993). In this regard, instead of recursively updating the upper triangular matrix $\mathbf{R}(k)$ as in QR decomposition recursive least squares, this technique recursively update the inverse of the upper triangular matrix $\mathbf{R}^{-1}(k)$ via sequence of Givens rotations. The foundation to develop this technique (IQRD-RLS) is based on the solution to triangular system of equation,

$$\mathbf{w}(k) = \mathbf{R}^{-1}(k)\mathbf{q}(k),$$

the vector $\mathbf{q}(k)$ is the transformed desired signal vector. The uniqueness of these techniques is that both algorithms irrespective of conditioning of the autocorrelation matrix converges with low misadjustment and capable of tracking incoming signals at a steady rate. This technique like direct QRD-RLS is numerically stable.

Orthogonal transform are preferred primarily because of their norm preserving property. Suppose for the 2- norm we have

$$\|\mathbf{Qv}\|_2^2 = (\mathbf{Qv})^T(\mathbf{Qv}) = \mathbf{v}^T\mathbf{Q}^T\mathbf{Qv} = \mathbf{v}^T\mathbf{v} = \|\mathbf{v}\|_2^2,$$

assuming that round off error is introduce in $\mathbf{v}$, then the error after an orthogonal transformation $\mathbf{Q}$ is

$$\|\mathbf{Q}\hat{\mathbf{v}}\|_2^2 = (\mathbf{Q}\hat{\mathbf{v}})^T(\mathbf{Q}\hat{\mathbf{v}}) = \hat{\mathbf{v}}^T\mathbf{Q}^T\mathbf{Q}\hat{\mathbf{v}} = \hat{\mathbf{v}}^T\hat{\mathbf{v}} = \|\hat{\mathbf{v}}\|_2^2.$$

Hence, no additional error is introduced by the orthogonal transformation. QR technique provides means of transforming data matrix into a simpler form while providing stable method for computation.

Based on discussions in section 3.4 we analyze the following. If small changes in the autocorrelation matrix cause little changes in the output data then the solution is well condition and stable. However, if small changes in the autocorrelation matrix cause large changes in the output data then the solution is ill-conditioned and unstable. We illustrate the above analysis using equation (1.1.1). If the autocorrelation matrix in (1.1.1) is ill-conditioned, the computed solution will not be exact. Otherwise if the autocorrelation matrix is well conditioned and the right pivoting technique is applied, one can compute exact solution. Suppose that the entries of the autocorrelation matrix are represented as floating point numbers, small roundoff errors may often occur during the reduction process which may affect the computed solution. The exact solution depends on the conditioning of the autocorrelation matrix. Suppose that we can measure the conditioning of the autocorrelation matrix, this process will be used to obtain a bound for the relative error in the computed solution. If the condition number is near to one, then the relative error and the relative residual error will be close, otherwise if the condition number is greater than one this mean that the relative error may be many times larger than the relative residual error (Leon S. J. 2006). Stability and instability are related to the algorithm and conditioning is a feature of the problem under consideration. A well conditioned system is stable whereas an ill-conditioned system is unstable. The condition number of the autocorrelation matrix of the problem play vital role in the convergence of these algorithms. The conditioning of the autocorrelation matrix determines the convergence behavior of these algorithms. The correlation parameter α is used to measure the conditioning of the autocorrelation matrix for the algorithms presented. From our experiment, as the correlation parameter varies the convergence of the RLS varies and variation in QR techniques is insignificant compared to RLS. This shows that the variation of the correlation parameter does not totally affect the convergence behavior of the QR techniques (Figures 7.3-7.4). However, variations in the correlation parameters drastically affect the convergence of RLS algorithm (Figure 7.2). These algorithms are evaluated based on rate of convergence and misadjustment. The system we identified is the finite duration impulse response transversal filter coefficient. The simulation is performed using the mean weight error norm to study the comparative performance of these algorithms. This approach provides better means of analyzing the rate of convergence, tracking and misadjustment. For low value of correlation parameter RLS converges fast but as the correlation parameter increases RLS show evidence of reduced tractability and produces high misadjustment (Figure 7.2). On the other hand, QRD-RLS and IQRD-RLS converge slow but as the correlation parameter increases misadjustment remain stable. The slow rate of convergence is because they are able to track incoming signal at a steady rate (Figure 7.3 and Figure 7.4). This thesis analyze the comparative performance and stability of the QR decomposition based algorithms for recursive least squares and recursive least squares algorithm for adaptive filtering applications.

1.2 Objectives of study

1) This thesis is designed to study recursive least squares (RLS), QR decomposition recursive least squares (QRD-RLS) and inverse QR decomposition recursive least squares (IQRD-RLS) algorithms based on Givens rotations.

2) To study and analyze the comparative performance of the recursive least squares (RLS), QR decomposition recursive least squares (QRD-RLS) and inverse QR decomposition recursive least squares algorithms using mean weight error norm (MWEN) to evaluate the convergence rate and misadjustment.

3) The stability of these algorithms are investigated based on simulation setup using system identification application. Simulation results show that QR decomposition based algorithms using Givens rotation is numerically stable than the recursive least squares using matrix inversion lemma.

1.3 Organization of the Thesis

The rest of this thesis is organized as follows. Literature review is presented in chapter two. The following are discussed to enhance our understanding of the concept under discussion: Introduction to adaptive filtering, adaptive filter structure, mathematical formulation of adaptive filtering problem and classification of adaptive filters are presented in chapter three. Recursive least squares and its derivation are described in chapter four and in chapter five, we prescribe step by step approach required to triangularized data matrix to upper triangular matrix via sequence of Givens rotations and updating of desired signal vector. This chapter includes the following concepts: QR technique for solving least squares problem, modification of QR technique for adaptive filtering problem, Givens rotations, Givens rotations recursive least squares for adaptive filtering, QR decomposition recursive least squares and its implementation is presented. Chapter six is a direct consequence of chapter five in the sense that the transformation process performed is extended to this chapter. This chapter includes the following: application of IQRD-RLS for adaptive filtering and updating parameters for computing coefficient vector of IQRD-RLS. In chapter seven, we perform simulation to evaluate the comparative performance of these techniques (RLS, QRD-RLS and IQRD-RLS), these algorithms are evaluated based on the mean weight error norm to analyze the rate of convergence and misadjustment. Conclusions and future work are presented in chapter eight.

Chapter 2

Literature Review

Recursive least squares is known to converge fast with high computational complexity and is proved to be numerically unstable (Alexander S.T. 1993) due to loss of positive definiteness, in order to resolve this problems researchers resorted to develop various techniques to improve numerical stability and reduce computational cost.

Orthogonal transformation (QR decomposition) using Givens rotations method was applied to triangularized data matrix into a set of simultaneous linear equation (Alexander S.T. 1993) which enable the coefficient vector to be computed via backward substitution. It was noted that this approach using QR decomposition is computationally expensive. To resolve this problem, researchers began to develop other techniques to resolve this impediment that is associated with this approach hence inverse QR decomposition was proposed and this approach does not require backward substitution to compute the coefficient vector (Alexander S.T. 1993; Diniz P. S. R. 2002; Apolinario (Jr.) J. A. 2009).

It was presented that it is very possible for RLS to become numerically unstable, (McWhiter J. G. 1994) applied QR decomposition to compute solutions to least squares problem and proved that it can be implemented efficiently in systolic array. In their presentation, they further suggest using Cholesky factor $\tilde{\mathbf{R}}$ instead of data autocorrelation matrix that inverse update techniques can be realize which utilizes orthogonal transformation in contrast to the former approach and explained further that inverse update can be applied to compute coefficient vector. McWhiter J. G. (1983) and Ward C. R., Hargrave P. J. and McWhiter J. G. (1986)

proposed adaptive filtering application for recursive least squares, this process was performed via Givens rotations to triangularize data matrix and implement it using systolic array. Their choice to apply it to systolic array was to avoid solving triangular system of equation in order to compute coefficient vector (de Campos M. L. R. 2009).

RLS based on Givens rotations using look ahead technique was applied to improve accuracy of redial basis function model (RBF). In computing RBF model coefficient using least squares technique numerical instability is often envisaged due to matrix inversion as such QR decomposition was introduced to eliminate numerical instability problems in the system (Shing T. 2002).

Order recursiveness using geometric interpretation of QR and inverse QR decomposition in least squares sense was investigated by (Apley W. D. 1995). They further suggest that order recursiveness is in track with the coefficient vector, error vector and residual error are in complete order of least squares projection as well as lower order least squares projection. They also investigated the geometric interpretation of orthogonal matrix and inverse Cholesky factor. QR decomposition based on recursive least squares using Givens rotations was applied to estimate memoryless complex polynomial coefficient which characterized high power amplifier. It was also used to investigate inverse polynomial model to achieve high power linearization, using QRD-RLS was aimed at achieving good numerical properties (Muruganathan S. D. 2006)

General linearly constrained recursive least squares problem using inverse QRD was proposed (Shiunn-Jang C. 2001) and applied to minimum variance filtering problems using Givens rotations to evaluate adaptation gain, the outcome of this process was to compute coefficient vector without backward substitution.

It was mentioned in (Wang S. 2008) that QRD-RLS converges faster than stochastic gradient descent approach. They investigated techniques to adjust coefficient vector to minimize constant modulus algorithm cost function (CMA), since the cost function is exponentially weighted sum their desire was to compute coefficient vector iteratively. They simply applied QRD-RLS techniques so that the system could be numerically robust, stable and easily implemented on pipeline structure based on the above mention they (Wang S. 2008) were able to develop QRD-RLS-CMA..

Adaptive algorithm based on recursive QR decomposition for the data matrix was proposed by (Athanasios A. R. 1996). They described that Givens rotations and modified Gram-Schmidt technique can provide modeling error directly without computing the coefficient vector. They however noted that these techniques are computationally complex if coefficient vector is required which means that backward substitution is required to compute coefficient vector. They applied Householder transformation which updates a square factor instead of triangular system and compute coefficient vector without backward substitution.

In this thesis, among various orthogonal techniques we apply Givens rotations due to its iterative nature since we aim at developing cost function that will be recursively updated. QR decomposition via Givens rotations is a notable adaptive filtering algorithm. We investigate the idea behind triangularization of the data matrix via QR decomposition using Givens rotations and different techniques applied to compute coefficient vector and associated errors (a posteriori and a priori errors). We apply recursive approach to recursively compute coefficient vector and associated errors in order to correspond to adaptive filtering application.

Chapter 3

Fundamentals of Adaptive Filtering and Adaptive Least Squares Problem

This chapter is the building block to understand and explain adaptive filtering and adaptive least squares problem. In doing this, we give detail analysis of adaptive filtering describing it via its applications, structures, algorithms and classifications. The problem we examine in this chapter is to formulate the cost function required for adaptive filtering which is minimized by adaptive algorithm in order to compute the coefficient vector $\mathbf{w}(k)$. The cost function $\varepsilon(k)$ consist of input signal $x(k)$, desired signal $d(k)$ and adaptive filter output $y(k)$, the cost function is formed via error signal. We further formulate least squares problem as an adaptive and exponentially weighted least squares problem.

3.1 Introduction to Adaptive Filter

Adaptive filters are self designing system which can adjust themselves to different environments. Filtering is a process of noise removal from a measured process in order to reveal or enhance information about some quantity of interest. Desired signal is generated from the FIR system (Diniz P.S.R. 2002). The desired signal may have introduce some noise during thermal or other physical effects related to signal generation system or it may introduce noise due to measuring system or a digital data sampling process. The type of application is defined by the choice of the signals acquired from the environment to be the input, desired and output signals. The following are some adaptive signal processing applications

1. System Identification

2. Adaptive Equalizer
3. Speech Coding
4. Adaptive Spectrum Analysis
5. Adaptive Noise Cancellation
6. Adaptive Beamforming
7. Echo Cancellation
8. Signal Enhancement

The common features of these applications that categorize them in this unique family is that they all involve a process of filtering some input signal to correspond to the desired response. Adaptive filter structure is shown in Figure 3.1; the parameters of adaptive filter can easily be obtain in Figure 3.1. The filter parameters are updated by making a set of measurements of the underlying signals and applying that set of signals to the adaptive filtering algorithm.

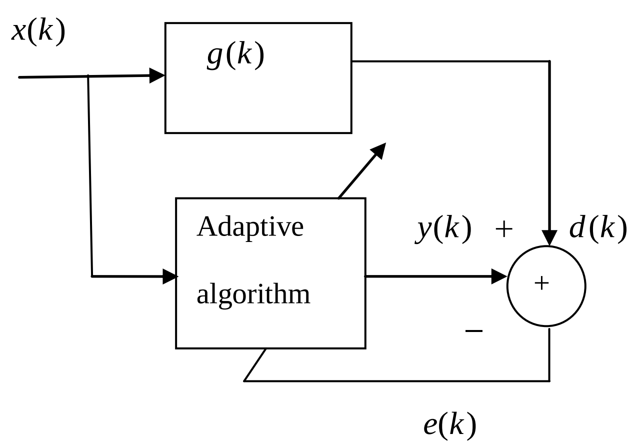

Figure 3.1 Adaptive filter configuration.

k : is the time index,

$x(k)$: denote the input signal,

$y(k)$: denote adaptive filter output,

$d(k)$: denote desired signal,

$\mathbf{w}(k)$: is adaptive tap weight coefficient vector,

$e(k) = d(k) - y(k)$: is the error signal and

$g(k)$: is the transfer function or plant.

Error signal is applied to adaptation algorithm to update the adaptive filter coefficient vector. The adaptation process aimed at minimizing some metric of the error signal by comparing the filter output to approximate the desired signal vector in a statistical sense (Sergio N. L. 2009). The signals applied to adaptation algorithm are the input signal and desired signal and error signal. If the signals are not well define the designed procedure is to model the signals and consequently design the filter. This procedure could be very expensive and difficult to implement online (Farhang - Boroujeny B. 1999). The solution to this problem is to employ an adaptive filter that performs online updating of its parameters through a rather simple algorithm using information available in the environment. In other word, adaptive filter performs data driven approximation procedure.

3.1.1 Adaptive Filter Structure

Adaptive filter structure can be implemented in different realizations. The choice of the filter structure can be influenced by computational complexity of the process and also the necessary number of iterations to achieve a desired performance level. There are two types of adaptive filter structures distinguished by the form of impulse response:

1. **Finite Duration Impulse Response Filter:** The most structure is the transversal filter Figure 3.2, which implements all zero transfer function with a direct form realization without feedback.

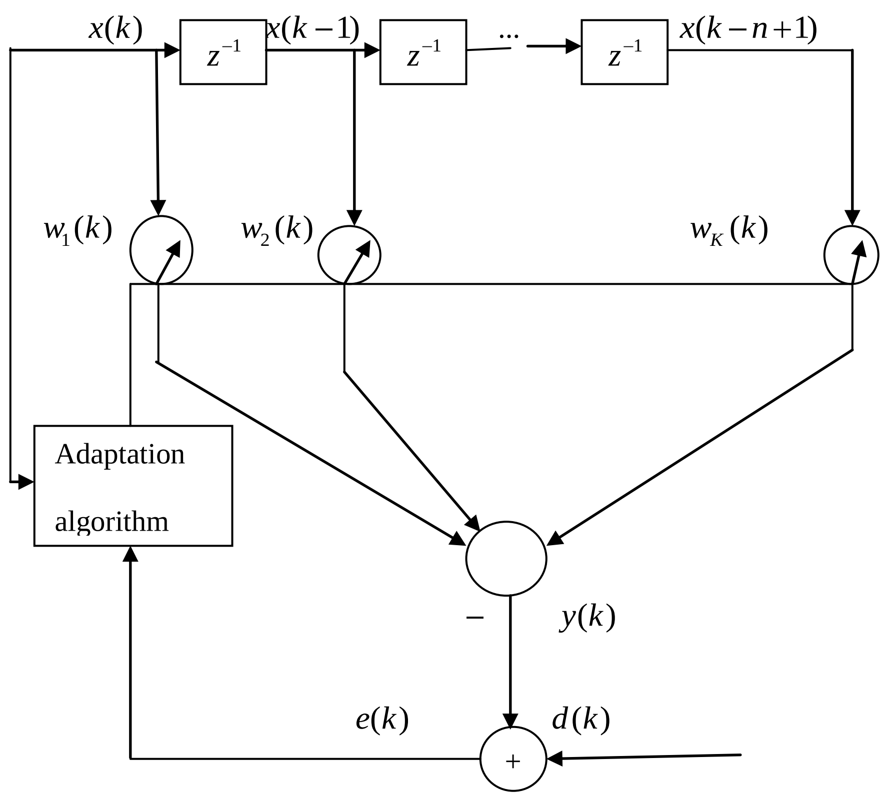

Figure 3.2 Transversal filter (FIR).

The output signal

$$y(k) = w_1(k)x(k) + w_2(k)x(k-1) + \ldots + w_K(k)x(k-n+1)$$

$$= \sum_{i=1}^{n} w_i(k)x(k-i+1) = \mathbf{w}^T(k)\mathbf{x}(k), \qquad (3.1.1.1)$$

$$\mathbf{x}(k) = \left[x(k) \quad \ldots \quad x(k-n+1) \right]^T,$$

is the instantaneous input signal vector.

$$\mathbf{w}(k) = \left[w_1(k) \quad \ldots \quad w_n(k) \right]^T,$$

is the approximate value for the weight vector after the kth system update, n is the filter order. We have different FIR structure designed to obtain improved structure compared to transversal filter structure in terms of computational complexity, speed of convergence and finite word length. Transversal filter has single input $x(k)$, filter output $y(k)$, and desired signal $d(k)$. Adaptive filter output is generated as a linear combination of the delayed sample of the input sample and the weight vector. The vector $\mathbf{w}(k)$ is the adaptive weight coefficient and $x(k-i)$ is the samples referred as filter tap input, the weight is controlled by the adaptation algorithm.

2) Infinite duration Impulse Response (IIR) filter: The most widely used realization of the adaptive IIR filter is the direct form realization due to simple implementation and analysis. There are inherent problems associated with this technique to recursive adaptive filters which are structure dependent for example pole stability monitoring requirement and slow speed of convergence. To circumvent these problems different realization were proposed to overcome the demerits mentioned above. An infinite duration impulse response (IIR) filter is governed by the recursive equation

$$\hat{y}(k) = \sum_{i=0}^{K-1} a_i(k)x(k-i) + \sum_{i=1}^{N-1} b_i(k)\hat{y}(k-i),$$

where $a_i(k)$ and $b_i(k)$ are the forward and feedback tap weights. This structure can easily become unstable since their poles may get shifted out of the unit circle by the adaptation process (Farhang - Boroujeny B. 1999). The cost function for the infinite duration impulse response filter has many local minimum points unlike FIR filter that has single global minimum point.

3.2 Mathematical Formulation of Adaptive Filtering Problem

In this section, we formulate the cost function applied in the rest of this thesis; the cost function is formulated based on finite impulse response (FIR) filter Figure 3.2. We have earlier defined adaptive filter output signal as a linear combination of the coefficient vector and the input vector which consist of the instantaneous input signal vector $\mathbf{x}(k)$ and approximate value of the weight vector. The idea is to formulate the cost function by defining the error vector which is minimized to obtain the coefficient vector. The cost function consist of desired signal $d(k)$ and adaptive filter output signal $y(k).$ Define $k \times n$ data matrix $\mathbf{X}(k)$ as

$$\mathbf{X}(k) = \begin{bmatrix} \mathbf{x}(1)^T \\ \mathbf{x}(2)^T \\ \vdots \\ \mathbf{x}(k)^T \end{bmatrix} = \begin{bmatrix} x(1) & 0 & \cdots & \cdots & 0 \\ x(2) & x(1) & 0 & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ x(k) & x(k-1) & \ddots & \vdots & x(1) \\ x(k+1) & x(k) & & & x(2) \\ \vdots & \vdots & & & \vdots \\ x(k) & x(k-1) & \cdots & \cdots & x(k-n+1) \end{bmatrix}. \qquad (3.2.1)$$

As new data is received into the system the data matrix becomes $\mathbf{X}(k+1)$ with dimension $(k+1) \times n,$ additional row $\mathbf{x}^T(k+1)$ is appended to the data matrix which can be expressed as

$$\mathbf{X}(k+1) = \begin{bmatrix} \mathbf{X}(k) \\ \cdots\ \cdots\ \cdots \\ \mathbf{x}^T(k+1) \end{bmatrix}.$$

The error signal is define as the difference between the desired signal $d(k)$ and adaptive filter output $y(k)$. The error signal $e(k)$ for the transversal FIR filter Figure 3.2 generated at the kth system update is define as

$$e(k) = d(k) - y(k) = d(k) - \mathbf{w}^T(k)\mathbf{x}(k). \tag{3.2.2}$$

The error signal vector computed after k system updates is

$$\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{X}(k)\mathbf{w}^T(k). \tag{3.2.3}$$

The vectors $\mathbf{e}(k)$ and $\mathbf{d}(k)$ are given respectively as

$$\mathbf{e}(k) = \begin{bmatrix} e(1) & e(2) & \cdots & e(k) \end{bmatrix}^T$$

and

$$\mathbf{d}(k) = \begin{bmatrix} d(1) & d(2) & \cdots & d(k) \end{bmatrix}^T.$$