In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

3/17/65
b

PROJECTIVE ITERATIVE SCHEMES FOR SOLVING SYSTEMS OF LINEAR EQUATIONS

A THESIS

Presented to

The Faculty of the Graduate Division

by

John Benjamin Hawkins

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Applied Mathematics

Georgia Institute of Technology

September, 1966

PROJECTIVE ITERATIVE SCHEMES FOR

SOLVING SYSTEMS OF LINEAR EQUATIONS

Approved:

Chairman

Date approved by Chairman: Aug 10, 1966

ACKNOWLEDGMENTS

I would like to express my sincere appreciation and thanks to Dr. William J. Kammerer for suggesting the topic and for his excellent help and suggestions in preparing and improving the thesis. I would also like to thank Dr. G. C. Caldwell and Dr. D. L. Finn for reading the paper and offering many helpful suggestions for its improvement. In addition, thanks are due to the National Science Foundation for the traineeship which I was awarded for the 1965-66 school year. I would also like to express my sincere thanks and appreciation to Mrs. Peggy Weldon for her excellent job of typing the thesis.

TABLE OF CONTENTS

INTRODUCTION

In many applications of mathematics to practical problems, it is necessary to determine the solution of a linear system of equations. One basic approach toward solving this type problem is to convert the problem to one of finding a fixed point of an equivalent system by some iterative method.

A solution to a linear system, say $A\bar{x} = \bar{b}$, where A is an nxn non-singular matrix, will be a fixed point of the equation

$$\bar{x} = \bar{x} + C_t(\bar{b} - A\bar{x}) \ . \tag{1}$$

This latter equation can be made into the iterative scheme

$$\bar{x}_{t+1} = \bar{x}_t + C_t(\bar{b} - A\bar{x}_t) \ , \quad t \geq 0 \ , \tag{2}$$

which will converge to a fixed point of (1) and the solution of $A\bar{x} = \bar{b}$ if sufficient conditions are placed on the C_t's. This paper will present a discussion of a special type of such iterative methods called projective iterative methods.

In Chapter I the definition of a projective iterative scheme is given, examples of projective schemes are discussed, and the question of when an iterative scheme is a projective method is answered. In Chapter II the questions of equivalence of an iterative scheme to the original system and convergence of the iterative scheme are discussed. Finally in Chapter III properties of a particular class of projective iterative schemes, known as norm reducing schemes, are investigated.

CHAPTER I

INTRODUCTION TO PROJECTIVE ITERATIVE SCHEMES

Here basic definitions and properties will be covered, examples of projective iterative schemes will be given, and the question of what schemes can be thought of as projective iterative schemes will be discussed.

Definition 1.1. The iterative scheme

$$\bar{x}_{t+1} = (I - C_t A)\bar{x}_t + C_t\bar{b} \;, \quad t \geq 0 \;, \tag{1.1}$$

is called a projective iterative method if, and only if, each $P_t = I - C_t A$, for $t \geq 0$, is a projection.

Properties of Projections

Because many of the results established later in this paper require the use of several basic properties of projections, it will facilitate this study if these properties are established prior to the main development of projective iterative methods.

Theorem 1.1: If P is a projection, $P \neq 0$, and $\|\cdot\|$ is a matrix norm, then $\|P\| \geq 1$.

Proof: For a non-zero projection P, the spectral radius is equal to one, $r_\sigma(P) = 1$. For a matrix norm,

$$\|P\| \geq r_\sigma(P) \;.$$

Hence, $r_\sigma(P) = 1 \leq \|P\|$.

Theorem 1.2: The matrix $P = I - CA$, where A is a non-singular matrix, is a projection if, and only if, $CAC = C$.

Proof: First it is assumed that P is a projection. Then $PP = P$. Therefore,

$$(I - CA)(I - CA) = (I - CA)$$

or

$$I - CA - CA + CACA = I - CA .$$

Hence,

$$CACA = CA$$

But A is non-singular, so $CAC = C$.

Now suppose $CAC = C$.

$$PP = (I - CA)(I - CA) = I - CA - CA + CACA$$

But $CAC = C$, so $CACA = CA$. Thus

$$PP = I - CA - CA + CA = I - CA = P .$$

Hence, P is a projection.

Theorem 1.3: Let G be a positive definite nxn matrix and Y be a nxk matrix, where k is a fixed integer satisfying $1 \leq k \leq n$, and the columns of Y are linearly independent. Then Y^*GY is non-singular.

Proof: It suffices to show Y^*GY is positive definite since this implies Y^*GY is non-singular. Let $\overline{x}$ be a k-dimensional vector.

Then

$$\bar{x}^*(Y^*GY)\bar{x} = (Y\bar{x})^*G(Y\bar{x}) \ .$$

Since G is positive definite,

$$(Y\bar{x})^*G(Y\bar{x}) \geq 0$$

with equality holding if, and only if, $Y\bar{x} = \bar{0}$. Now $Y\bar{x}$ is a linear combination of the columns of Y and hence can be the zero vector if, and only if, $\bar{x} = \bar{0}$ since the columns of Y are linearly independent. Thus

$$\bar{x}^*(Y^*GY)\bar{x} \geq 0$$

for all $\bar{x}$ with equality if, and only if, $\bar{x} = \bar{0}$. Therefore Y^*GY is positive definite and hence non-singular.

<u>Theorem 1.4</u>: Let Y be an nxk matrix, $1 \leq k \leq n$, with linearly independent columns. Let G be a positive definite matrix. Then

$$P = (I - Y(Y^*GY)^{-1}Y^*G)$$

is a projection.

Proof: From Theorem 1.3 $(Y^*GY)^{-1}$ exists and hence P is well defined.

$$PP = (I - Y(Y^*GY)^{-1}Y^*G)(I - Y(Y^*GY)^{-1}Y^*G)$$

$$PP = I - 2Y(Y^*GY)^{-1}Y^*G + Y(Y^*GY)^{-1}Y^*GY(Y^*GY)^{-1}Y^*G$$

Now $(Y^*GY)^{-1}Y^*GY = I$. Hence

$$PP = I - 2Y(Y^*GY)^{-1}Y^*G + YI(Y^*GY)^{-1}Y^*G ,$$

or

$$PP = I - Y(Y^*GY)^{-1}Y^*G = P .$$

Hence P is a projection.

Examples of Projective Iterative Schemes

In order for the reader to become better acquainted with the idea of projective iterative methods, several examples of iterative schemes for solving $A\bar{x} = \bar{b}$ which can be thought of in this manner are now cited.

I. Gauss-Seidel Iterative Scheme

Under this scheme the t^{th} iteration is given by

$$\bar{x}_t = \bar{x}_{t-1} + C_{t-1}(\bar{b} - A\bar{x}_{t-1}) \tag{1.2}$$

where, for $1 \leq t \leq n$,

$$C_{t-1} = \left[\begin{array}{c:c:c} O_{t-1} & 0 & 0 \\ \hdashline 0 & \frac{1}{a_{tt}} & 0 \\ \hdashline 0 & 0 & O_{n-t} \end{array}\right],$$

O_k denotes the $k \times k$ zero matrix, and for $t \geq n + 1$,

$$C_{t-1} = C_{(t-1)\text{mod } n}$$

($(t-1)$ mod n being used to denote the remainder when $t-1$ is divided by n).

Theorem 1.5: The Gauss-Seidel iterative scheme (1.2) is a projective iterative scheme.

Proof: It need only be shown that $I - C_tA$ is a projection for $t \geq 0$. From the fact that the C_t's are repeated and from Theorem 1.2, it is sufficient to show that $C_{t-1}AC_{t-1} = C_{t-1}$ for $1 \leq t \leq n$.

$$C_{t-1}A = \left[\begin{array}{c|c|c} 0_{t-1} & 0 & 0 \\ \hline 0 & \frac{1}{a_{tt}} & 0 \\ \hline 0 & 0 & 0_{n-t} \end{array}\right] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{t1} & \cdots & a_{tn} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$$C_{t-1}A = \left[\begin{array}{c} 0 \\ \hline \frac{a_{t1}}{a_{tt}} \cdots \frac{a_{tt}}{a_{tt}} \cdots \frac{a_{tn}}{a_{tt}} \\ \hline 0 \end{array}\right]$$

$$C_{t-1}AC_{t-1} = \left[\begin{array}{c} 0 \\ \hline \frac{a_{t1}}{a_{tt}} \cdots \frac{a_{tt}}{a_{tt}} \cdots \frac{a_{tn}}{a_{tt}} \\ \hline 0 \end{array}\right] \left[\begin{array}{c|c|c} 0_{t-1} & 0 & 0 \\ \hline 0 & \frac{1}{a_{tt}} & 0 \\ \hline 0 & 0 & 0_{n-t} \end{array}\right]$$

$$C_{t-1}AC_{t-1} = \left[\begin{array}{c} 0 \\ \hline 0 \cdots 0 \quad \frac{a_{tt}}{a_{tt}a_{tt}} \quad 0 \cdots 0 \\ \hline 0 \end{array}\right] = C_{t-1}$$

Thus by Theorem 1.2, $I - C_{t-1}A$ is a projection for $1 \leq t \leq n$. Hence $I - C_tA$ is a projection for all $t \geq 0$. Thus the Gauss-Seidel iterative scheme is a projective iterative scheme.

II. Block Gauss-Seidel Iterative Scheme

For the system $A\bar{x} = \bar{b}$, let the matrix A be partitioned in the form

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix},$$

where each A_{ii}, $1 \leq i \leq N$, is a square non-singular matrix. Let v_i be the number of rows A_{ii} has for each i, $1 \leq i \leq N$, and let

$$p_i = \sum_{j=1}^{i} v_j .$$

Then the $(t+1)^{st}$ iteration for the block Gauss-Seidel scheme is given by

(1.3) $$\bar{x}_{t+1} = \bar{x}_t + C_t(\bar{b} - A\bar{x}_t), \quad t \geq 0 ,$$

where for $0 \leq t \leq N-1$,

$$C_t = \left[\begin{array}{c:c:c} 0_{p_t} & 0 & 0 \\ \hdashline 0 & A^{-1}_{t+1\,t+1} & 0 \\ \hdashline 0 & 0 & 0_{n-p_{t+1}} \end{array}\right],$$

and for $t \geq N$,

$$C_t = C_{t \bmod N} .$$

<u>Theorem 1.6</u>: The block Gauss-Seidel iterative scheme (1.3) is a projective iterative scheme.

Proof: It need only be shown that $I - C_tA$ is a projection for each $t \geq 0$. From the fact that the C_t's repeat and from Theorem 1.2, it need only be shown that

$$C_{t-1}AC_{t-1} = C_{t-1}$$

for $1 \leq t \leq N$.

$$C_{t-1}A = \left[\begin{array}{c:c:c} O_{p_{t-1}} & 0 & 0 \\ \hdashline 0 & A_{tt}^{-1} & 0 \\ \hdashline 0 & 0 & O_{N-p_t} \end{array}\right] \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{t1} & \cdots & A_{tN} \\ \vdots & & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix}$$

$$C_{t-1}A = \left[\begin{array}{ccccc} & & 0 & & \\ \hdashline A_{tt}^{-1}A_{t1} & \cdots & A_{tt}^{-1}A_{tt} & \cdots & A_{tt}^{-1}A_{tN} \\ \hdashline & & 0 & & \end{array}\right]$$

$$C_{t-1}AC_{t-1} = \left[\begin{array}{ccccc} & & 0 & & \\ \hdashline A_{tt}^{-1}A_{t1} & \cdots & A_{tt}^{-1}A_{tt} & \cdots & A_{tt}^{-1}A_{tN} \\ \hdashline & & 0 & & \end{array}\right] \left[\begin{array}{c:c:c} P_{t-1} & 0 & 0 \\ \hdashline 0 & A_{tt}^{-1} & 0 \\ \hdashline 0 & 0 & O_{N-p_t} \end{array}\right]$$

$$C_{t-1}AC_{t-1} = \begin{bmatrix} & & 0 & & \\ \hline 0 \cdots 0 & & A_{tt}^{-1}A_{tt}A_{tt}^{-1} & & 0 \cdots 0 \\ \hline & & 0 & & \end{bmatrix} = C_{t-1}$$

Thus $I - C_{t-1}A$ is a projection for $1 \leq t \leq N$, and since the C_t's repeat, the block Gauss-Seidel iterative scheme is a projective iterative scheme.

III. Method of Steepest Descent

Here it is assumed that the matrix A is real, positive definite, and symmetric. Under these assumptions

$$\|\bar{x}\|_A = (\bar{x}^* A \bar{x})^{1/2}$$

defines a vector norm.

Now if one denotes the solution of the system $A\bar{x} = \bar{b}$ by $\bar{u}$ and defines $\bar{s} = \bar{u} - \bar{x}$ where $\bar{x}$ is any vector, the problem of finding the solution of $A\bar{x} = \bar{b}$ can be considered as the problem of finding the vector $\bar{x}$ for which

$$g(\bar{x}) = \|\bar{s}\|_A^2$$

is minimized (This holds since $\|\bar{s}\|_A^2 \geq 0$ and equality holds only when $\bar{s} = \bar{0}$, i.e. when $\bar{x} = \bar{u}$).

Now suppose $\bar{x}_{p-1}$ is an approximation to the minimum point of g and hence the solution of $A\bar{x} = \bar{b}$. In the method of steepest descent the next approximation to the solution is obtained by traveling along a line through $\bar{x}_{p-1}$ pointing in the direction in which g changes most rapidly at $\bar{x}_{p-1}$ until a point is reached at which the distance from the solution to the point on this line in terms of $\|\cdot\|_A$ is a minimum. This point is then taken as the next approximation to the solution.

From elementary calculus one has that g changes most rapidly at a point in the direction of the gradient vector of g at that point. This vector is given by

$$\overline{\text{Del}}\ g = \begin{pmatrix} \frac{\partial g}{\partial x_1} \\ \vdots \\ \frac{\partial g}{\partial x_n} \end{pmatrix} .$$

Now for each i, $1 \leq i \leq n$,

$$\frac{\partial g(\bar{x})}{\partial x_i} = \frac{\partial}{\partial x_i} \left((\bar{u} - \bar{x})^* A (\bar{u} - \bar{x}) \right)$$

$$\frac{\partial g(\bar{x})}{\partial x_i} = - \frac{\partial \bar{x}^*}{\partial x_i} A(\bar{u} - \bar{x}) + (\bar{u} - \bar{x})^* A\left(- \frac{\partial \bar{x}}{\partial x_i}\right) .$$

Note that

$$\frac{\partial \bar{x}^*}{\partial x_i} A(\bar{u} - \bar{x}) = (\bar{u} - \bar{x})^* A^* \left(\frac{\partial \bar{x}}{\partial x_i}\right) ;$$

but $A = A^*$ so

$$\frac{\partial \bar{x}^*}{\partial x_i} A(\bar{u} - \bar{x}) = (\bar{u} - \bar{x})^* A\left(\frac{\partial \bar{x}}{\partial x_i}\right) .$$

Thus

$$\frac{\partial g(\bar{x})}{\partial x_i} = -2\left(\frac{\partial \bar{x}}{\partial x_i}\right)^* A(\bar{u} - \bar{x}) \ .$$

Now $\frac{\partial \bar{x}}{\partial x_i} = \bar{e}_i$ where $\bar{e}_i$ has all zero elements except for the i^{th} element which is one. Therefore

$$\frac{\partial g(\bar{x})}{\partial x_i} = -2\bar{e}_i^* A(\bar{u} - \bar{x}) = -2\bar{e}_i^*(A\bar{u} - A\bar{x}) \ .$$

But $A\bar{u} = \bar{b}$, so

$$\frac{\partial g(\bar{x})}{\partial x_i} = -2\bar{e}_i^*(\bar{b} - A\bar{x}) = -2\bar{e}_i^*\bar{r}(\bar{x})$$

where $\bar{r}(\bar{x}) = \bar{b} - A\bar{x}$, the residual vector at $\bar{x}$. Note that the i^{th} row of the identity matrix, I, is given by $\bar{e}_i^*$ so that

$$\bar{r}(\bar{x}) = I\bar{r}(\bar{x}) = \begin{bmatrix} \bar{e}_1^* \\ \vdots \\ \bar{e}_i^* \\ \vdots \\ \bar{e}_n^* \end{bmatrix} \bar{r}(\bar{x}) = \begin{bmatrix} \bar{e}_1^*\bar{r}(\bar{x}) \\ \\ \bar{e}_i^*\bar{r}(\bar{x}) \\ \\ \bar{e}_n^*\bar{r}(\bar{x}) \end{bmatrix}$$

Hence,

$$\overline{\text{Del}}\ g(\bar{x}) = \begin{bmatrix} \frac{\partial g(\bar{x})}{\partial x_1} \\ \vdots \\ \frac{\partial g(\bar{x})}{\partial x_i} \\ \vdots \\ \frac{\partial g(\bar{x})}{\partial x_n} \end{bmatrix} = -2 \begin{bmatrix} \bar{e}_1^*\bar{r}(\bar{x}) \\ \\ \bar{e}_i^*\bar{r}(\bar{x}) \\ \\ \bar{e}_n^*\bar{r}(\bar{x}) \end{bmatrix} = -2\bar{r}(\bar{x}) \ .$$

Thus g changes most rapidly at $\bar{x}$ in the direction of $\bar{r}(\bar{x})$. Hence to get the next approximation to the solution after $\bar{x}_{p-1}$, one travels on the line through $\bar{x}_{p-1}$ in the direction of $\bar{r}_{p-1} = \bar{b} - A\bar{x}_{p-1}$ until the distance, in terms of the norm, $\|\circ\|_A$, to $\bar{u}$ is minimized. Hence, one seeks the value of λ for which $g(\bar{x}_{p-1} + \lambda\, \bar{r}_{p-1})$ is minimized. Now $g(\bar{x}_{p-1} + \lambda\, \bar{r}_{p-1})$ will be put in a form which will make it easy to tell which value of λ minimizes it.

$$
\begin{aligned}
g(\bar{x}_{p-1} + \lambda\, \bar{r}_{p-1}) &= (\bar{u} - \bar{x}_{p-1} - \lambda\, \bar{r}_{p-1})^* A(\bar{u} - \bar{x}_{p-1} - \lambda\, \bar{r}_{p-1}) \\
&= (\bar{u} - \bar{x}_{p-1})^* A(\bar{u} - \bar{x}_{p-1}) - \lambda \bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1}) \\
&\quad - \lambda(\bar{u} - \bar{x}_{p-1})^* A\bar{r}_{p-1} + \lambda^2\, \bar{r}^*_{p-1} A\bar{r}_{p-1}
\end{aligned}
$$

$$
\begin{aligned}
g(\bar{x}_{p-1} + \lambda\bar{r}_{p-1}) &= \|\bar{u} - \bar{x}_{p-1}\|_A^2 - \|\bar{r}_{p-1}\|_A^2 \frac{(\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1}))^2}{\|\bar{r}_{p-1}\|_A^4} \\
&\quad + \|\bar{r}_{p-1}\|_A^2 \left[\lambda^2 - 2\lambda \frac{\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1})}{\|\bar{r}_{p-1}\|_A^2} + \frac{(\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1}))^2}{\|\bar{r}_{p-1}\|_A^4}\right]
\end{aligned}
$$

$$
\begin{aligned}
g(\bar{x}_{p-1} + \lambda\bar{r}_{p-1}) &= \|\bar{u} - \bar{x}_{p-1}\|_A^2 - \frac{\|\bar{r}_{p-1}\|_A^2\, (\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1}))^2}{\|\bar{r}_{p-1}\|_A^4} \\
&\quad + \|\bar{r}_{p-1}\|_A^2 \left[\lambda - \frac{\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1})}{\|\bar{r}_{p-1}\|_A^2}\right]^2
\end{aligned}
$$

Note that the only variable is λ and the term involving it is non-negative.

Hence, $g(\bar{x}_{p-1} + \lambda\, \bar{r}_{p-1})$ will be minimized if λ is chosen to make the term in which it appears zero. This will be the case if λ is chosen to be

$$\lambda = \frac{\bar{r}^*_{p-1} A(\bar{u} - \bar{x}_{p-1})}{\|\bar{r}_{p-1}\|_A^2} = \frac{\bar{r}^*_{p-1}(A\bar{u} - A\bar{x}_{p-1})}{\bar{r}^*_{p-1} A\bar{r}_{p-1}}$$

But $\bar{r}_{p-1} = \bar{b} - A\bar{x}_{p-1} = A\bar{u} - A\bar{x}_{p-1}$. Thus

$$\lambda = \frac{\bar{r}^*_{p-1}\bar{r}_{p-1}}{\bar{r}^*_{p-1} A\bar{r}_{p-1}} .$$

Thus the next approximation to the solution is given by

$$\bar{x}_p = \bar{x}_{p-1} + \lambda\, \bar{r}_{p-1} = \bar{x}_{p-1} + \frac{\bar{r}^*_{p-1}\bar{r}_{p-1}}{\bar{r}^*_{p-1} A\bar{r}_{p-1}}\, \bar{r}_{p-1} .$$

Hence the t^{th} iterate in the method of steepest descent is given by

$$(1.4) \qquad \bar{x}_t = \bar{x}_{t-1} + \frac{\bar{r}^*_{t-1}\bar{r}_{t-1}}{\bar{r}^*_{t-1} A\bar{r}_{t-1}}\, \bar{r}_{t-1}, \quad t \geq 1 ,$$

where $\bar{r}_{t-1} = \bar{b} - A\bar{x}_{t-1}$.

<u>Theorem 1.7</u>: The method of steepest descent (1.4) is a projective iterative scheme.

Proof: For each $t \geq 0$,

$$\bar{x}_{t+1} = \bar{x}_t + \frac{\bar{r}^*_t\bar{r}_t}{\bar{r}^*_t A\bar{r}_t}\, \bar{r}_t = \bar{x}_t + \frac{\bar{r}^*_t(\bar{b} - A\bar{x}_t)}{\bar{r}^*_t A\bar{r}_t}\, \bar{r}_t$$

$$\bar{x}_{t+1} = \left[I - \bar{r}_t \frac{\bar{r}_t^* A}{\bar{r}_t^* A \bar{r}_t} \right] \bar{x}_t + \bar{r}_t \frac{\bar{r}_t^* \bar{b}}{\bar{r}_t^* A \bar{r}_t}$$

Letting $C_t = \bar{r}_t \dfrac{\bar{r}_t^*}{\bar{r}_t^* A \bar{r}_t}$ gives

$$\bar{x}_{t+1} = (I - C_t A)\bar{x}_t + C_t \bar{b} \ .$$

Letting $G = A$ and $Y = \bar{r}_t$ in Theorem 1.4 gives, using the result of that theorem, that $I - C_t A$ is a projection, for each $t \geq 0$. Hence, the method of steepest descent is a projective iterative scheme.

IV. Conjugate Direction Iterative Schcmes

Another group of iterative schemes which are projective iterative schemes is the class of conjugate direction schemes. These have the property that they terminate within n iterations with the exact solution of the system $A\bar{x} = \bar{b}$.

Basically these methods fall into the following general scheme. Matrices C and B are selected so that the matrix $R = CAB$ is a positive definite matrix. Then a set of linearly independent vectors, $\bar{v}_1, \ldots, \bar{v}_n$, is chosen, either in advance or as the iterations are computed, so that the vectors, $\bar{v}_i$, are R-orthogonal, i.e. $\bar{v}_i^* R \bar{v}_j = 0$ if $i \neq j$, $1 \leq i, j \leq n$. An initial guess to the solution, say $\bar{x}_0$, is chosen and succeeding iterates are chosen as follows. Assuming that $\bar{x}_k$ has been found, $k < n$, $\bar{x}_{k+1}$ is found by proceeding along the line through $\bar{x}_k$ in the direction of $\bar{v}_{k+1}$ until the point is reached which minimizes the distance to the true solution in the sense of the norm generated by R. This point is then taken as the next iterate. Since the

$\bar{v}_i$'s are R-orthogonal, the new iterate also has the property that it is the closest point in the R-norm sense to the actual solution on the lines in the directions of the previously used $\bar{v}_i$'s passing through that point. Because the $\bar{v}_i$'s span the space, the n^{th} iterate, $\bar{x}_n$, gives the solution. In terms of an algebraic equation the $(t+1)^{st}$ iterate is given by

$$\bar{x}_{t+1} = \bar{x}_t + \frac{\bar{v}^*_{t+1} C \bar{r}_t B \bar{v}_{t+1}}{\bar{v}^*_{t+1} R \bar{v}_{t+1}} , \quad 0 \leq t \leq n-1 , \tag{1.5}$$

where $\bar{r}_t = \bar{b} - A\bar{x}_t$.

It is shown by Faddeev and Faddeeva [1] and for the case when $B = I$ by Hestenes and Stiefel [3] that $\bar{x}_n$ is indeed the solution of $A\bar{x} = \bar{b}$.

<u>Theorem 1.8</u>: The iterative scheme (1.5) is a projective iterative scheme.

Proof: For $0 \leq t \leq n-1$, (1.5) can be written

$$\bar{x}_{t+1} = \bar{x}_t + \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1} C(\bar{b} - A\bar{x}_t)}{\bar{v}^*_{t+1} R \bar{v}_{t+1}} .$$

$$\bar{x}_{t+1} = \left[I - \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1} CA}{\bar{v}^*_{t+1} R \bar{v}_{t+1}} \right] \bar{x}_t + \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1} C\bar{b}}{\bar{v}^*_{t+1} R \bar{v}_{t+1}}$$

Thus the above is a projective iterative scheme if

$$I - \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1} CA}{\bar{v}^*_{t+1} R \bar{v}_{t+1}}$$

is a projection for each t, $0 \le t \le n-1$. Now by Theorem 1.2 this will be true if

$$\frac{B\bar{v}_{t+1}\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}} \cdot A \cdot \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}} = \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}},$$

Recalling the fact that $R = CAB$, one has

$$\frac{B\bar{v}_{t+1}\bar{v}^*_{t+1}CAB\bar{v}_{t+1}\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}\;\bar{v}^*_{t+1}R\bar{v}_{t+1}} = \frac{B\bar{v}_{t+1}(\bar{v}^*_{t+1}R\bar{v}_{t+1})\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}\bar{v}^*_{t+1}R\bar{v}_{t+1}}$$

$$= \frac{B\bar{v}_{t+1}\bar{v}^*_{t+1}C}{\bar{v}^*_{t+1}R\bar{v}_{t+1}}.$$

Thus (1.5) is a projective iterative scheme.

What Schemes are Projective Schemes?

The question now arises as to just exactly what iterative schemes can be thought of in the context of projective iterative methods. The following series of theorems provides the answer.

Lemma 1.1. Let m be a positive integer greater than 1. For each w, $1 \le w \le m-1$, define R_w by

$$R_w = \left[\begin{array}{c|cc|c} I_{w-1} & \multicolumn{2}{c|}{0} & 0 \\ \hline & \lambda & 1 & \\ 0 & & & 0 \\ & 0 & 1 & \\ \hline 0 & \multicolumn{2}{c|}{0} & I_{m-w-1} \end{array}\right].$$

(The convention is made that when I_0 appears in the above representation, the column and row in which that symbol appears are omitted from the array.) Then

$$R_{m-1}R_{m-2}\cdots R_1 = \begin{bmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & \ddots & \ddots \\ 0 & & & \lambda & 1 \\ & & & & 1 \end{bmatrix}.$$

Proof: First, the following result is established using mathematical induction. For k such that $2 \leq k \leq m-1$,

$$R_kR_{k-1}\cdots R_1 = \left[\begin{array}{ccc|c|c} \lambda & 1 & 0 & 0 & \\ & \ddots & \ddots & \vdots & 0 \\ 0 & \lambda & 1 & 0 & \\ & & \lambda & 1 & \\ \hline & 0 & & 1 & 0 \\ \hline & 0 & & 0 & I_{m-k-1} \end{array}\right].$$

Let $k = 2$.

$$R_2R_1 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & 1 & \\ \hline & 0 & & I_{m-3} \end{array}\right] \left[\begin{array}{ccc|c} \lambda & 1 & 0 & \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \\ \hline & 0 & & I_{m-3} \end{array}\right] = \left[\begin{array}{ccc|c} \lambda & 1 & 0 & \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & 1 & \\ \hline & 0 & & I_{m-3} \end{array}\right].$$

Hence the statement holds for $k = 2$. Now assume the statement is true for $k \leq m-2$ and consider the product $R_{k+1}R_k \cdots R_1$. Using the induction hypothesis,

$$R_{k+1}R_k\cdots R_1 = \left[\begin{array}{c|c|c} I_k & 0 & 0 \\ \hline 0 & \begin{matrix} \lambda & 1 \\ 0 & 1 \end{matrix} & 0 \\ \hline 0 & 0 & I_{m-k-2} \end{array}\right] \left[\begin{array}{c|c|c} \begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & \lambda \end{matrix} & \begin{matrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & 0 \\ \hline 0 & 0 & I_{m-k-2} \end{array}\right].$$

$$R_{k+1}R_k\cdots R_1 = \left[\begin{array}{c|c|c} \begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{matrix} & \begin{matrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} \lambda & 1 \\ 0 & 1 \end{matrix} & 0 \\ \hline 0 & 0 & I_{m-k-2} \end{array}\right].$$

Hence if the statement holds for $k \leq m-2$ it holds for $k+1$. This completes the induction and the statement holds. Now taking $k = m-1$,

$$R_{m-1}R_{m-2}\cdots R_1 = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & 1 \end{bmatrix}.$$

Hence the lemma is established.

<u>Lemma 1.2</u>: For each i, $1 \leq i \leq k$, let K_i be an nxn matrix which can

be partitioned in the form

$$K_i = \left[\begin{array}{c|cccc} 0 & 0 & \cdots & & 0 \\ \hline 0 & B_1 & & & 0 \\ \vdots & & B_2 & & \\ & & & \ddots & \\ 0 & & 0 & & B_k \end{array}\right],$$

where B_j is a $t_j \times t_j$ matrix given either by

$$B_j = I_{t_j}$$

or by

$$B_j = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

The matrix B_i has the second of the above possible representations. Let $\overline{K}_i$ be the matrix K_i with B_i changed to I_{t_i}. Then K_i can be written as a product of projections times $\overline{K}_i$.

Proof: Let $p = \sum_{j=1}^{i-1} t_j$ and $q = \sum_{j=i+1}^{k} t_j$.

Case I. $B_i = [\lambda]$. Then $t_i = 1$.

Let

$$P_i^1 = \left[\begin{array}{c|c|c|c} 0 & 0 & \lambda & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right].$$

Note that $P_i^1 P_i^1 = P_i^1$. Therefore P_i^1 is a projection. Let

$$P_i^2 = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right].$$

Note that $P_i^2 P_i^2 = P_i^2$. Therefore P_i^2 is a projection. **Let**

$$P_i^3 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right].$$

Note that $P_i^3 P_i^3 = P_i^3$. Therefore P_i^3 is a projection.

Claim: $P_i^3 P_i^2 P_i^1 \bar{K}_i = K_i$.

$$P_i^3 P_i^2 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right] \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3 P_i^2 P_i^1 = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c|c|c|c} 0 & 0 & \lambda & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3P_i^2P_i^1\overline{K}_i = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & \begin{matrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_{i-1} \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & \begin{matrix} B_{i+1} & & 0 \\ & \ddots & \\ 0 & & B_k \end{matrix} \end{array}\right]$$

$$P_i^3P_i^2P_i^1\overline{K}_i = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & \begin{matrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_{i-1} \end{matrix} & 0 & 0 \\ \hline 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & \begin{matrix} B_{i+1} & & 0 \\ & \ddots & \\ 0 & & B_k \end{matrix} \end{array}\right] = K_i$$

Thus K_i can be written as a product of projections times $\overline{K}_i$.

Case II.

$$B_i = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & 0 \\ & & \ddots & \ddots \\ & 0 & \lambda & 1 \\ & & & \lambda \end{bmatrix}.$$

Let R_w be defined as in Lemma 1.1, for each w such that $1 \leq w \leq t_i-1$. Then

$$R_{t_i-1}R_{t_i-2}\cdots R_1 = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & 1 \end{bmatrix}.$$

For each w such that $1 \leq w \leq t_i - 1$, define F_w by

$$F_w = \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & R_w & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right].$$

Then

$$F_{t_i-1}F_{t_i-2}\cdots F_1 = \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & R_{t_i-1}\cdots R_1 & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right]$$

$$F_{t_i-1}F_{t_i-2}\cdots F_1 = \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & \begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & 1 \end{matrix} & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right].$$

Let

$$P_i^1 = \left[\begin{array}{c:c:c:c:c} 0 & 0 & 0 & \lambda & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 \\ \hdashline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_q \end{array}\right],$$

Note that $P_i^1 P_i^1 = P_i^1$. Thus P_i^1 is a projection. Let

$$P_i^2 = \left[\begin{array}{c|c|c|c|c} 1 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right].$$

Note that $P_i^2 P_i^2 = P_i^2$. Thus P_i^2 is a projection. Let

$$P_i^3 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right].$$

Note that $P_i^3 P_i^3 = P_i^3$. Thus P_i^3 is a projection. It will now be shown that

$$P_i^3 P_i^2 P_i^1 F_{t_i-1} \cdots F_1 = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & B_i & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right].$$

$$P_i^3 P_i^2 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right] \left[\begin{array}{c|c|c|c|c} 1 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3 P_i^2 = \left[\begin{array}{c|c|c|c|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3 P_i^2 P_i^1 = \left[\begin{array}{c|c|c|c|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c|c|c|c|c} 0 & 0 & 0 & \lambda & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3 P_i^2 P_i^1 = \left[\begin{array}{c|c|c|c|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 \\ \hline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda & 0 \\ \hline 0 & 0 & 0 & 0 & I_q \end{array}\right].$$

Now recalling the expression previously derived for $F_{t_i-1} \cdots F_1$, one has the following.

$$P_i^3P_i^2P_i^1F_{t_i-1}\cdots F_1 = \left[\begin{array}{c:c:c:c:c} 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 \\ \hdashline 0 & 0 & I_{t_i-1} & 0 & 0 \\ \hdashline 0 & 0 & 0 & \lambda & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c:c:c:c:c} 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 \\ \hdashline 0 & 0 & \begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & \lambda \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$P_i^3P_i^2P_i^1F_{t_i-1}\cdots F_1 = \left[\begin{array}{c:c:c:c:c} 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 \\ \hdashline 0 & 0 & \begin{matrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & \lambda \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} & 0 \\ \hdashline 0 & 0 & 0 & \lambda & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

Hence,

$$(1.6) \qquad P_i^3P_i^2P_i^1F_{t_i-1}\cdots F_1 = \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & B_i & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right]$$

It will now be shown that for each w, $1 \leq w \leq t_i-1$, F_w can be written as a product of five projections. Note that this establishes that the matrix in equation (1.6) can be written as a product of projections.

since P_i^1, P_i^2, and P_i^3 are projections.

Let

$$U_w^1 = \left[\begin{array}{c:c:c:c:c:c} 0 & 0 & 0 & \lambda & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_{t_i-w} & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right].$$

Then $U_w^1 U_w^1 = U_w^1$. Hence U_w^1 is a projection.

Let

$$U_w^2 = \left[\begin{array}{c:c:c:c:c:c} 1 & 0 & 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right].$$

Then $U_w^2 U_w^2 = U_w^2$. Hence U_w^2 is a projection.

Let

$$U_w^3 = \left[\begin{array}{c:c:c:cc:c:c}
1 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hdashline
0 & I_p & 0 & \multicolumn{2}{c:}{0} & 0 & 0 \\ \hdashline
0 & 0 & I_{w-1} & \multicolumn{2}{c:}{0} & 0 & 0 \\ \hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & 0 & 1 & & \\ \hdashline
0 & 0 & 0 & \multicolumn{2}{c:}{0} & I_{t_i-w-1} & 0 \\ \hdashline
0 & 0 & 0 & \multicolumn{2}{c:}{0} & 0 & I_q
\end{array}\right].$$

Then $U_w^3 U_w^3 = U_w^3$. Hence U_w^3 is a projection.

Let

$$U_w^4 = \left[\begin{array}{c:c:c:cc:c:c}
1 & 0 & 0 & \multicolumn{2}{c:}{0} & 0 & 0 \\ \hdashline
0 & I_p & 0 & \multicolumn{2}{c:}{0} & 0 & 0 \\ \hdashline
0 & 0 & I_{w-1} & \multicolumn{2}{c:}{0} & 0 & 0 \\ \hdashline
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & & & 0 & 1 & & \\ \hdashline
0 & 0 & 0 & \multicolumn{2}{c:}{0} & I_{t_i-w-1} & 0 \\ \hdashline
0 & 0 & 0 & \multicolumn{2}{c:}{0} & 0 & I_q
\end{array}\right].$$

Then $U_w^4 U_w^4 = U_w^4$. Hence U_w^4 is a projection.

Let

$$U_w^5 = \left[\begin{array}{c:c}
0 & 0 \\ \hdashline
0 & I_{n-1}
\end{array}\right].$$

Then $U_w^5 U_w^5 = U_w^5$. Hence U_w^5 is a projection.

Now

$$U_w^5 U_w^4 = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right] \left[\begin{array}{c|c|c|c|c|c} 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline \begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-1-w} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right].$$

$$U_w^5 U_w^4 = \left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline \begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$U_w^5 U_w^4 U_w^3 = \left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline \begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c|c|c|c|c|c} 1 & 0 & 0 & 1 \quad 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

$$
U_w^5 U_w^4 U_w^3 = \left[\begin{array}{c|c|c|c|c|c}
0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & I_p & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline
\begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & I_q
\end{array}\right]
$$

$U_w^5 U_w^4 U_w^3 U_w^2 =$

$$
\left[\begin{array}{c|c|c|c|c|c}
0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & I_p & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline
\begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & I_q
\end{array}\right]
\left[\begin{array}{c|c|c|c|c|c}
1 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & I_p & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & I_q
\end{array}\right]
$$

$$
U_w^5 U_w^4 U_w^3 U_w^2 = \left[\begin{array}{c|c|c|c|c|c}
0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & I_p & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline
\begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline
0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & I_q
\end{array}\right]
$$

$U_w^5 U_w^4 U_w^3 U_w^2 U_w^1 =$

$$\left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline \begin{matrix} 1 \\ 0 \end{matrix} & 0 & 0 & \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & \lambda \quad 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right]$$

Hence,

$$U_w^5 U_w^4 U_w^3 U_w^2 U_w^1 = \left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_{w-1} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & \begin{matrix} \lambda & 1 \\ 0 & 1 \end{matrix} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{t_i-w-1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_q \end{array}\right] = F_w .$$

Thus F_w can be written as a product of five projections.

Hence, using the same notation as above, equation (1.6) becomes

$$P_i^3 P_i^2 P_i^1 F_{t_i-1} \cdots F_1 = P_i^3 P_i^2 P_i^1 \prod_{w=1}^{t_i-1} (U_w^5 U_w^4 U_w^3 U_w^2 U_w^1)$$

or

$$P_i^3 P_i^2 P_i^1 \prod_{w=1}^{t_i-1} (U_w^5 U_w^4 U_w^3 U_w^2 U_w^1) = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & B_i & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] .$$

Now

$$K_i = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & I_p & 0 & 0 \\ \hline 0 & 0 & B_i & 0 \\ \hline 0 & 0 & 0 & I_q \end{array}\right] \bar{K}_i ,$$

because

$$K_i = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & \begin{matrix} B_1 & & \\ & \ddots & \\ & & B_k \end{matrix} \end{array}\right] \text{ and } \bar{K}_i = \left[\begin{array}{c|c|c|c} 0 & 0 & 0 & 0 \\ \hline 0 & \begin{matrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_{i-1} \end{matrix} & 0 & 0 \\ \hline 0 & 0 & I_{t_i} & 0 \\ \hline 0 & 0 & & \begin{matrix} B_{i+1} & & 0 \\ & \ddots & \\ 0 & & B_k \end{matrix} \end{array}\right],$$

so that

$$\left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & B_i & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right] \overline{K}_i = \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & B_i & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right] \left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & \begin{matrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_{i-1} \end{matrix} & 0 & 0 \\ \hdashline 0 & 0 & I_{t_i} & 0 \\ \hdashline 0 & 0 & 0 & \begin{matrix} B_{i+1} & & 0 \\ & \ddots & \\ 0 & & B_k \end{matrix} \end{array}\right]$$

$$= \left[\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & \begin{matrix} B_1 & & \\ & \ddots & \\ & & B_k \end{matrix} \end{array}\right] = K_i \ .$$

But it has been shown that

$$\left[\begin{array}{c:c:c:c} 0 & 0 & 0 & 0 \\ \hdashline 0 & I_p & 0 & 0 \\ \hdashline 0 & 0 & B_i & 0 \\ \hdashline 0 & 0 & 0 & I_q \end{array}\right]$$

can be written as a product of projections. Hence, K_i can be written as a product of projections times $\overline{K}_i$. This completes the proof of the lemma.

<u>Theorem 1.9</u>. Any singular matrix A can be written as a product of projections.

Proof: If A is a 1x1 matrix, then $A = [0]$, and hence A is a projection and the theorem holds.

Now suppose A is an nxn matrix for $n \geq 2$. There exists a non-singular matrix D so that $D^{-1}AD$ is the Jordan normal form of A where

$$D^{-1}AD = \begin{bmatrix} J_0 & & & 0 \\ & J_1 & & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix};$$

and J_i is a $t_i x t_i$ matrix, $0 \leq i \leq k$, $\sum_{i=0}^{k} t_i = n$;

$$J_0 = [0] \quad \text{or} \quad J_0 = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{bmatrix};$$

$$J_i = [\lambda] \quad \text{or} \quad J_i = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$$

where λ is an eigenvalue of A.

Case I. Suppose $J_0 = [0]$. Then $k > 0$. Note that in this case $D^{-1}AD$ is of the form of the matrix K_1 in Lemma 1.2. Hence applying Lemma 1.2, there exist projections $P^1_1, \ldots, P^1_{b_1}$ so that

$$D^{-1}AD = \prod_{i_1=1}^{b_1} P^1_{i_1} \begin{bmatrix} 0 & & & & \\ & I_{t_1} & & 0 & \\ & & J_2 & & \\ & 0 & & \ddots & \\ & & & & J_k \end{bmatrix}.$$

But now

$$\begin{bmatrix} 0 & & & & \\ & I_{t_1} & & 0 & \\ & & J_2 & & \\ & 0 & & \ddots & \\ & & & & J_k \end{bmatrix}$$

is a matrix of the form K_2 in Lemma 1.2, whence applying the lemma again, there exist projections $P_1^2, \ldots, P_{b_2}^2$ so that

$$\begin{bmatrix} 0 & & & & \\ & I_{t_1} & & 0 & \\ & & J_2 & & \\ & 0 & & \ddots & \\ & & & & J_k \end{bmatrix} = \prod_{i_2=1}^{b_2} P_{i_2}^2 \begin{bmatrix} 0 & & & & & \\ & I_{t_1} & & 0 & & \\ & & I_{t_2} & & & \\ & 0 & & J_3 & & \\ & & & & \ddots & \\ & & & & & J_k \end{bmatrix}.$$

Hence,

$$D^{-1}AD = \left(\prod_{i_1=1}^{b_1} P_{i_1}^1\right)\left(\prod_{i_2=1}^{b_2} P_{i_2}^2\right) \begin{bmatrix} 0 & & & & & \\ & I_{t_1} & & 0 & & \\ & & I_{t_2} & & & \\ & 0 & & J_3 & & \\ & & & & \ddots & \\ & & & & & J_k \end{bmatrix}.$$

Continuing to apply Lemma 1.2 in this manner yields

$$D^{-1}AD = \prod_{j=1}^{k} \prod_{i_j=1}^{b_j} P_{i_j}^j \begin{bmatrix} 0 & & & & \\ & I_{t_1} & & 0 & \\ & & I_{t_2} & & \\ & 0 & & \ddots & \\ & & & & I_{t_k} \end{bmatrix}$$

where now $P^j_{i_j}$ are the appropriate projections for each j and i_j such that $1 \leq i_j \leq b_j$ and $1 \leq j \leq k$. Now

$$\begin{bmatrix} 0 & & & \\ & I_{t_1} & & 0 \\ & & \ddots & \\ 0 & & & I_{t_k} \end{bmatrix} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & I_{n-1} \end{array}\right]$$

which is a projection. Denote it by P. Then

$$D^{-1}AD = \left(\prod_{j=1}^{k}\left(\prod_{i_j=1}^{b_j} P^j_{i_j}\right)\right)P$$

whence

$$A = D\left(\prod_{j=1}^{k}\left(\prod_{i_j=1}^{b_j} P^j_{i_j}\right)\right)PD^{-1} = \left(\prod_{j=1}^{k}\left(\prod_{i_j=1}^{b_j} DP^j_{i_j}D^{-1}\right)\right)DPD^{-1}.$$

But each of the terms of the product on the right is a projection since given any projection M and non-singular matrix N, NMN^{-1} is also a projection. Thus A can be expressed as a product of projections.

Case II. Suppose

$$J_0 = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{bmatrix},$$

where J_0 is a $t_0 \times t_0$ matrix $(t_0 \geq 2)$. First suppose $k > 0$. Thus

$$D^{-1}AD = \left[\begin{array}{c|c} \begin{matrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{matrix} & 0 \\ \hline 0 & \begin{matrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{matrix} \end{array}\right] = \left[\begin{array}{c|c|c} \begin{matrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \\ & & & 0 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{matrix} \end{array}\right]$$

By Case I the matrix

$$\begin{bmatrix} 0 & & & \\ & J_1 & & 0 \\ & & \ddots & \\ & 0 & & J_k \end{bmatrix}$$

can be written as a product of projections, say $P_1, \ldots, P_b$, i.e.,

$$P_b P_{b-1} \cdots P_1 = \begin{bmatrix} 0 & & 0 \\ & J_1 & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix}.$$

Let

$$F_i = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i \end{array}\right]$$

for each i, $1 \le i \le b$. Then

$$F_i F_i = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i \end{array}\right] \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i \end{array}\right] = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i P_i \end{array}\right]$$

$$F_i F_i = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i \end{array}\right] = F_i$$

since $P_i P_i = P_i$ as P_i is a projection. Thus F_i is a projection.

Consider

$$\prod_{i=1}^{b} F_i = \prod_{i=1}^{b} \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & P_i \end{array}\right]$$

$$\prod_{i=1}^{b} F_i = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & \prod_{i=1}^{b} P_i \end{array}\right] = \left[\begin{array}{c|c|c} I_{t_0-1} & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{matrix} \end{array}\right].$$

Now consider the matrix

$$J_0 = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{bmatrix}$$

Define the matrices R_w, $1 \leq w \leq t_0-1$, by

$$R_w = \left[\begin{array}{c|c|c} I_{w-1} & 0 & 0 \\ \hline 0 & \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} & 0 \\ \hline 0 & 0 & I_{t_0-w-1} \end{array}\right],$$

Then these are of the same form as the matrices defined in Lemma 1.1

with $\lambda = 0$. Thus the conclusion of this lemma holds and

$$R_{t_0-1}\cdots R_1 = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 1 \end{bmatrix}.$$

Note that $R_w R_w = R_w$ for $1 \leq w \leq t_0 - 1$. Thus each R_w is a projection. Now let R be defined by

$$R = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & 0 \end{array}\right].$$

Then R is a projection and

$$RR_{t_0-1}\cdots R_1 = \left[\begin{array}{c|c} I_{t_0-1} & 0 \\ \hline 0 & 0 \end{array}\right] \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & 0 & 1 \\ 0 & & & 0 \end{bmatrix} = J_0.$$

Thus J_0 can be written as a product of projections. Let G_w for $1 \leq w \leq t_0-1$ be defined by

$$G_w = \left[\begin{array}{c|c} R_w & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right].$$

Then each G_w is a projection since $G_w G_w = G_w$. Define G by

$$G = \left[\begin{array}{c|c} R & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right].$$

Then G is also a projection since $GG = G$. Now note that

$$G \prod_{w=1}^{t_0-1} G_w = \left[\begin{array}{c|c} R & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right] \prod_{w=1}^{t_0-1} \left(\left[\begin{array}{c|c} R_w & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right] \right)$$

$$= \left[\begin{array}{c|c} R \displaystyle\prod_{w=1}^{t_0-1} R_w & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right] = \left[\begin{array}{c|c} J_0 & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right].$$

Now recalling what $\prod_{i=1}^{b} F_i$ is equal to, one has

$$\left(\prod_{i=1}^{b} F_i\right)(G)\left(\prod_{w=1}^{t_0-1} G_w\right) = \left[\begin{array}{c|c|c} I_{t_0-1} & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} J_1 & 0 \\ 0 & J_K \end{matrix} \end{array}\right] \left[\begin{array}{c|c} J_0 & 0 \\ \hline 0 & I_{n-t_0} \end{array}\right]$$

$$= \left[\begin{array}{c|c|c} I_{t_0-1} & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} J_1 & 0 \\ 0 & J_K \end{matrix} \end{array}\right] \left[\begin{array}{ccc|c} 0 & 1 & 0 & \\ & & 0 & 0 \\ 0 & 0 \quad 1 & 0 & \\ & 0 & & \\ \hline & 0 & 0 & 0 \\ \hline & 0 & 0 & I_{n-t_0} \end{array}\right]$$

$$= \left[\begin{array}{c|c|c} \begin{matrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & 0 & 1 \\ & & & 0 \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & \begin{matrix} J_1 & & 0 \\ & \ddots & \\ 0 & & J_k \end{matrix} \end{array}\right]$$

$$\left(\prod_{i=1}^{b} F_i\right)(G)\left(\prod_{w=1}^{t_0-1} G_w\right) = \begin{bmatrix} J_0 & & & 0 \\ & J_1 & & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix} = D^{-1}AD \ ,$$

Therefore $D^{-1}AD$ can be written as a product of projections. Now

$$A = D\left(\prod_{i=1}^{b} F_i\right)(G)\left(\prod_{w=1}^{t_0-1} G_w\right)D^{-1}$$

or

$$A = \left(\prod_{i=1}^{b} DF_iD^{-1}\right)(DGD^{-1})\left(\prod_{w=1}^{t_0-1} DG_wD^{-1}\right) .$$

But DF_iD^{-1}, $1 \le i \le b$, DGD^{-1}, DG_wD^{-1}, $1 \le w \le t_0-1$, are projections since F_i, $1 \le i \le b$, G, and G_w, $1 \le w \le t_0 - 1$, are projections. Therefore A can be written as a product of projections.

Now suppose $D^{-1}AD = J_0$. In the proof of the preceding part it was shown that J_0 can be written as a product of projections. Denote them by $G_1,\ldots,G_r$ where

$$G_rG_{r-1}\ldots G_1 = J_0 = D^{-1}AD \ .$$

Therefore

$$A = D\left(\prod_{i=1}^{r} G_i\right) D^{-1} = \prod_{i=1}^{r} DG_iD^{-1} .$$

But DG_iD^{-1}, $1 \leq i \leq r$, are projections since each G_i is a projection. Thus A can be written as a product of projections. This completes the proof.

Corollary 1.9.1. An iterative scheme,

$$\bar{x}_{t+1} = (I - C_tA)\bar{x}_t + C_t\bar{b} , \quad t \geq 0 ,$$

can be written in the form of a projective iterative scheme if for each $t \geq 0$, $I - C_tA = I$ or $I - C_tA$ is singular.

Proof: Let $t \geq 0$ be given and consider the corresponding matrix $I - C_tA$. If this is singular, then by Theorem 1.9 projections $P_0^t, \ldots, P_{r_t}^t$ exist such that

$$I - C_tA = \prod_{i=0}^{r_t} P_i^t$$

and thus C_t can be represented as

$$C_t = (I - \prod_{i=0}^{r_t} P_i^t) A^{-1} .$$

Now let

$$\bar{y}_{i+1}^t = P_i^t \bar{y}_i^t + (I - P_i^t) A^{-1} \bar{b}$$

for $0 \le i \le r_t$, where $\bar{y}_0^t = \bar{x}_t$.

It is now shown that, for $1 \le j \le r_t + 1$,

$$\bar{y}_j^t = P_{j-1}^t \cdots P_0^t \bar{y}_0^t + (I - P_{j-1}^t \cdots P_0^t) A^{-1} \bar{b} .$$

This is established by induction. The case for $j = 1$ is obvious from the definition. It is now assumed that the conclusion holds for some positive integer $j < r_t + 1$. Then for $j + 1$ one has

$$\bar{y}_{j+1}^t = P_j^t \bar{y}_j^t + (I - P_j^t) A^{-1} \bar{b} .$$

Using the induction hypothesis gives

$$\bar{y}_{j+1}^t = P_j^t (P_{j-1}^t \cdots P_0^t \bar{y}_0^t + (I - P_{j-1}^t \cdots P_0^t) A^{-1} \bar{b}) + (I - P_j^t) A^{-1} \bar{b} .$$

Thus

$$\bar{y}_{j+1}^t = P_j^t \cdots P_0^t \bar{y}_0^t + P_j^t A^{-1} \bar{b} - P_j^t \cdots P_0^t A^{-1} \bar{b} + A^{-1} \bar{b} - P_j^t A^{-1} \bar{b}$$

or

$$\bar{y}_{j+1}^t = P_j^t \cdots P_0^t \bar{y}_0^t + (I - P_j^t \cdots P_0^t) A^{-1} \bar{b} .$$

Hence, the induction is complete, and the claim established. Taking $j = r_t + 1$ gives

$$\bar{y}_{r_t+1}^t = P_{r_t}^t \cdots P_0^t \bar{y}_0^t + (I - P_{r_t}^t \cdots P_0^t) A^{-1} \bar{b} .$$

But

$$I - C_t A = P^t_{r_t} \cdots P^t_0 \quad \text{and} \quad \bar{y}^t_0 = \bar{x}_t \; .$$

Therefore

$$\bar{y}^t_{r_t+1} = (I - C_t A)\bar{x}_t + (I - I + C_t A)A^{-1}\bar{b}$$

or

$$\bar{y}^t_{r_t+1} = (I - C_t A)\bar{x}_t + C_t\bar{b} = \bar{x}_{t+1} \; .$$

If $I - C_t A = I$, $C_t A = 0$ so $C_t = 0$ and $\bar{x}_{t+1} = \bar{x}_t$.

Let $r_t = 0$ and $\bar{y}^t_{r_t+1} = \bar{x}_t = \bar{x}_{t+1}$.

Consider the scheme

$$\bar{y}^t_{i+1} = P^t_i \bar{y}^t_i + (I - P^t_i)A^{-1}\bar{b}, \quad 0 \le i \le r_t, \quad t \ge 0,$$

where $\bar{y}^t_0 = \bar{y}^{t-1}_{r_{t-1}+1}$.

Then this is a projective iterative scheme which agrees with the original scheme for each $i = r_t$, $t \ge 0$, that is,

$$\bar{y}^t_{r_t+1} = \bar{x}_{t+1} \; .$$

Thus the original scheme can be written as a projective iterative scheme.

Before one jumps to the conclusion that all iterative schemes can be thought of as projective iterative schemes, it should be noted that there are a great many which cannot. Indeed consider the schemes of the form

$$\bar{x}_{t+1} = (I - C_t A)\bar{x}_t + C_t\bar{b}, \quad t \geq 0 ,$$

where $I - C_tA$ is non-singular and $I - C_tA \neq I$. If this scheme were equivalent to a projective iterative scheme, then each $I - C_tA$ would be a projection or else be expressible as a (finite) product of projections. Either of these cases requires that $I - C_tA$ be the identity or singular. Since these possibilities have been ruled out, the above scheme is not equivalent to a projective scheme.

CHAPTER II

EQUIVALENCE AND CONVERGENCE OF SCHEME

The question now arises as to just exactly what projective iterative schemes actually give rise to solutions of the original system. Here questions of equivalence of the iterative scheme to the original system and convergence will be considered.

Equivalence to Original System

Sufficient conditions must be placed on the $P_t = I - C_t A$, $t \geq 0$, in iteration (1.1) so that a fixed point of (1.1) will be a solution of the original system, $A\bar{x} = \bar{b}$. The following theorem provides one such set of conditions.

Theorem 2.1: If, given any integer k, there exists a positive integer n_k such that one is not an eigenvalue of the matrix

$$B = \prod_{j=0}^{n_k} P_{k+j} ,$$

then a fixed point of (1.1) is the solution of $A\bar{x} = \bar{b}$.

Proof: Since one is not an eigenvalue of B, then if $\bar{x}$ is a vector such that

$$\bar{x} = B\bar{x} ,$$

then $\bar{x} = \bar{0}$. From (1.1)

$$\bar{x}_{k+n_k+1} = P_{k+n_k}\,\bar{x}_{k+n_k} + C_{n_k+k}\,\bar{b}\ .$$

Working backwards one gets

$$\bar{x}_{k+n_k+1} = \left[\prod_{j=0}^{n_k} P_{k+j}\right]\bar{x}_k + \sum_{j=0}^{n_k-1}\left[\prod_{i=j+1}^{n_k} P_{k+i}\right] C_{k+j}\bar{b} + C_{k+n_k}\bar{b}\ .$$

Suppose $\bar{y}$ is a fixed point of (1.1), i.e.

$$\bar{y} = P_t\bar{y} + C_t\bar{b}\ , \quad t \geq 0\ .$$

Then substituting into the above expression and recalling the definition of B gives

$$\bar{y} = B\bar{y} + \sum_{j=0}^{n_k-1}\left[\prod_{i=j+1}^{n_k} P_{k+i}\right] C_{k+j}\bar{b} + C_{k+n_k}\bar{b}\ .$$

Now the solution, $\bar{u}$, of $A\bar{x} = \bar{b}$ is also a fixed point of (1.1) and hence the above equation. Thus

$$\bar{u} = B\bar{u} + \sum_{j=0}^{n_k-1}\left[\prod_{i=j+1}^{n_k} P_{k+i}\right] C_{k+j}\bar{b} + C_{k+n_k}\bar{b}\ .$$

Hence subtracting the last two equations gives

$$\bar{u} - \bar{y} = B(\bar{u} - \bar{y}) + \bar{0}\ .$$

Therefore $\bar{u} - \bar{y} = \bar{0}$ or $\bar{u} = \bar{y}$ since one is not an eigenvalue of B. Thus a fixed point of (1.1) is the solution of $A\bar{x} = \bar{b}$.

Convergence

Note that the above discussion does not deal with the question of when an iterative scheme will converge, but only states that if an iterative scheme satisfies certain conditions, then the fixed point is the solution to the original system. Now the question of when a projective iterative scheme will converge will be taken up.

Theorem 2.2: Let P_t, $t \geq 0$, be the matrices associated with the iteration (1.1). Suppose there exists an $M \geq 1$ such that $\|P_t\| \leq M$ for all $t \geq 0$. Suppose further that there exist integers n and k_0 such that

$$\left\| \prod_{j=0}^{n-1} P_{k+j} \right\| \leq \alpha < 1$$

for all $k \geq k_0$. Then the iterative scheme (1.1) will converge to a solution of the system $A\bar{x} = \bar{b}$.

Proof: Let $\bar{u}$ be the solution of $A\bar{x} = \bar{b}$. From (1.1) one has

$$\bar{x}_{t+1} = P_t\bar{x}_t + C_t\bar{b}$$

where $P_t = I - C_tA$, $t \geq 0$. Since $\bar{u}$ is the solution of $A\bar{x} = \bar{b}$, $\bar{u}$ is a fixed point of (1.1) and hence

$$\bar{u} = P_t\bar{u} + C_t\bar{b}$$

for all $t \geq 0$. Let $\bar{s}_{t+1} = \bar{u} - \bar{x}_{t+1}$. Then

$$\bar{s}_{t+1} = \bar{u} - \bar{x}_{t+1} = P_t(\bar{u} - \bar{x}_t) + C_t\bar{b} - C_t\bar{b} \ .$$

Thus

$$\bar{s}_{t+1} = P_t \bar{s}_t ,$$

Working backwards one has

$$\bar{s}_{t+1} = \left[\prod_{j=0}^{t} P_j \right] \bar{s}_0 .$$

Thus

$$\|\bar{s}_{t+1}\| \leq \| \prod_{j=0}^{t} P_j \| \, \|\bar{s}_0\| .$$

It will now be shown that

$$\| \prod_{j=0}^{t} P_j \|$$

approaches zero as t tends to infinity. Taking $t > k_0$, let w be the largest integer so that $k_0 + wn \leq t$. Then one can write

$$\prod_{j=0}^{t} P_j = \left[\prod_{j=k_0+wn}^{t} P_j \right] \left[\prod_{m=0}^{w-1} \left[\prod_{i=k_0+mn}^{k_0+(n+1)m-1} P_i \right] \right] \left[\prod_{j=0}^{k_0-1} P_j \right] .$$

Now

$$\| \prod_{j=0}^{t} P_j \| \leq \left\| \prod_{j=k_0+wn}^{t} P_j \right\| \left\| \prod_{m=0}^{w-1} \left[\prod_{i=k_0+mn}^{k_0+(n+1)m-1} P_i \right] \right\| \left\| \prod_{j=0}^{k_0-1} P_j \right\| .$$

But

$$\left\| \prod_{j=k_0+wn}^{t} P_j \right\| \leq \prod_{j=k_0+wn}^{t} \|P_j\| \leq M^{t-k_0-wn+1} \leq M^n$$

since $M \geq 1$ and $k_0 + wn \leq t < k_0 + w(n+1)$. Also

$$\left\| \prod_{j=0}^{k_0-1} P_j \right\| \leq \prod_{j=0}^{k_0-1} \|P_j\| \leq M^{k_0} .$$

Thus

$$\left\| \prod_{j=0}^{t} P_j \right\| \leq M^{n+k_0} \left\| \prod_{m=0}^{w-1} \prod_{i=k_0+mn}^{k_0+(n+1)m-1} P_i \right\| .$$

Hence

$$\left\| \prod_{j=0}^{t} P_j \right\| \leq M^{n+k_0} \prod_{m=0}^{w-1} \left\| \prod_{i=k_0+mn}^{k_0+(n+1)m-1} P_i \right\| .$$

Now, using the hypothesis,

$$\left\| \prod_{i=k_0+mn}^{k_0+(n+1)m-1} P_i \right\| \leq \alpha < 1$$

for all $m \geq 0$. Thus

$$\left\| \prod_{j=0}^{t} P_j \right\| \leq M^{n+k_0} \alpha^w .$$

As t tends to infinity, w tends to infinity, and hence α^w approaches zero. Thus

$$0 \leq \lim_{t \to \infty} \left\| \prod_{j=0}^{t} P_j \right\| \leq \lim_{t \to \infty} M^{n+k_0} \alpha^w = 0 .$$

Therefore

$$\lim_{t\to\infty} \left\| \prod_{j=0}^{t} P_j \right\| = 0 .$$

Hence

$$\lim_{t\to\infty} \|\bar{s}_{t+1}\| = \lim_{t\to\infty} \left\| \prod_{j=0}^{t} P_j \right\| \|\bar{s}_0\| = 0 .$$

Thus

$$\lim_{t\to\infty} \bar{s}_{t+1} = \bar{0} ,$$

and hence

$$\bar{u} = \lim_{t\to\infty} \bar{x}_t ,$$

Thus the iteration (1.1) converges to the solution of the equation $A\bar{x} = \bar{b}$. This completes the proof.

Note that Theorem 2.2 gives only a sufficient condition for convergence. The hypothesis can be modified to give other conditions under which the iterative scheme (1.1) will converge. For example, the condition that there exist integers n and k_0 so that

$$\left\| \prod_{j=0}^{n-1} P_{k+j} \right\| \leq \alpha < 1, \quad k \geq k_0 ,$$

can be relaxed somewhat to the condition that there exist integers n and k_0 so that

$$\left\| \prod_{j=0}^{n-1} P_{k_0+mn+j} \right\| \leq \alpha < 1$$

for all integers $m \geq 0$. In fact this latter condition, which follows from the first, was actually all that was used in the proof. Other conditions can be given, but these only make matters seem more complicated and hence are omitted here.

The following series of theorems will be used to find conditions of the P_t which will indicate whether a particular iteration will converge.

Before stating the next theorem, it is necessary to define what is meant by an elliptic norm.

<u>Definition 2.1.</u> Let G be a positive definite Hermitian matrix. Then the norm defined by

$$\|\bar{x}\|_G = (\bar{x}^* G \bar{x})^{1/2}$$

is an elliptic norm. (It is easily verified that $\|\cdot\|_G$ is a norm.)

<u>Theorem 2.3:</u> If P is a non-zero projection on E_n, then there exists an elliptic norm such that $\|P\| = 1$, i.e., P is an orthogonal projection relative to some inner product.

Proof: Let M be the range of P and N be the null space of P. Let $(\cdot,\cdot)$ be the normal inner product associated with E_n, i.e., if $\bar{x}$, $\bar{y}$ are in E_n,

$$(\bar{x},\bar{y}) = \sum_{i=1}^{n} x_i y_i^*$$

where x_i is the i^{th} component of $\bar{x}$ and y_i^* is the complex conjugate of the i^{th} component of $\bar{y}$. Let $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_r$ be a basis for the range of P, and let $\bar{v}_{r+1}, \ldots, \bar{v}_n$ be a basis for the null space of P if N is not the set containing only the zero vector. From properties of projections the vector space E_n can be written as the direct sum of M and N. Thus the vectors $\bar{v}_1, \ldots, \bar{v}_n$ form a basis of E_n. Now let $\bar{x}$ be an element of E_n and let $c_1, \ldots, c_n$ be the coefficients in its expansion in terms of the $\bar{v}_i$, i.e.,

$$\bar{x} = \sum_{i=1}^{n} c_i \bar{v}_i .$$

Similarly let $\bar{y}$ be an element of E_n and let $d_1, \ldots, d_n$ be the coefficients in its expansion in terms of the $\bar{v}_i$. Define

$$(\bar{x}, \bar{y})_P = \sum_{i=1}^{n} c_i d_i^*$$

where d_i^* denotes the complex conjugate of d_i. It can be easily shown that this is an inner product.

It will now be shown that the norm arising from this inner product,

$$\|\bar{x}\|_P = [(\bar{x}, \bar{x})_P]^{1/2} ,$$

is an elliptic norm. Let $\bar{e}_i$, $1 \leq i \leq n$, be the usual basis for E_n consisting of the elementary unit vectors, i.e., each $\bar{e}_i$ has all zero elements except the i^{th} which is one. Now let

$$[a_{1i}, a_{2i}, \ldots, a_{ni}]^T$$

be the representation of $\bar{e}_i$ in terms of the $\bar{v}_i$'s, i.e.,

$$\bar{e}_i = \sum_{j=1}^{n} a_{ji}\bar{v}_j, \quad 1 \leq i \leq n .$$

Let $\bar{x}$ be an element in E_n with components x_i in terms of the $\bar{e}_i$'s. Then

$$\bar{x} = \sum_{i=1}^{n} x_i\bar{e}_i = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ji} \bar{v}_j .$$

Let A be the matrix (a_{ji}), i.e., the matrix whose columns are the representations of the $\bar{e}_i$'s in terms of the $\bar{v}_j$'s. Then in terms of the $\bar{v}_j$'s a vector $\bar{x}$ in E_n is represented by $A\bar{x}$. Thus $(.,.)_P$ gives rise to the elliptic norm defined by

$$\|\bar{x}\|_P = (\bar{x},\bar{x})_P^{1/2} = [\bar{x}^* A^* A\bar{x}]^{1/2} .$$

It will now be shown that P is an orthogonal projection relative to the inner product $(.,.)_P$. Let $\bar{x}$ be an element of E_n. Now $P\bar{x}$ is in the range of P so that

$$P\bar{x} = \sum_{i=1}^{n} b_i\bar{v}_i$$

where $b_i = 0$ for $r+1 \leq i \leq n$. But $\bar{x} - P\bar{x}$ is in the null space of P so that

$$\bar{x} - P\bar{x} = \sum_{i=1}^{n} c_i\bar{v}_i .$$

where $c_i = 0$ for $1 \leq i \leq r$. Thus

$$(\bar{x} - P\bar{x},\ P\bar{x})_P = \sum_{i=1}^{n} c_i b_i^* = 0$$

since for all i, $1 \leq i \leq n$, either $c_i = 0$ or $b_i = 0$. Hence P is an orthogonal projection with respect to this inner product.

It will now be shown that $\|P\|_P = 1$. It is already known that $\|P\|_P \geq 1$, so it need only be shown that $\|P\|_P \leq 1$. Let $\bar{x}$ be an element of E_n.

$$\|\bar{x}\|_P^2 = (\bar{x} + P\bar{x} - P\bar{x},\ \bar{x} + P\bar{x} - P\bar{x})_P$$

$$\|\bar{x}\|_P^2 = (P\bar{x},\ \bar{x} - P\bar{x} + P\bar{x})_P + (\bar{x} - P\bar{x},\ \bar{x} - P\bar{x} + P\bar{x})_P$$

$$\|\bar{x}\|_P^2 = (P\bar{x},\ P\bar{x})_P + (P\bar{x},\ \bar{x} - P\bar{x})_P + (\bar{x} - P\bar{x},\ P\bar{x})_P + (\bar{x} - P\bar{x},\ \bar{x} - P\bar{x})_P$$

$$\|\bar{x}\|_P^2 = (P\bar{x},\ P\bar{x})_P + (\bar{x} - P\bar{x},\ \bar{x} - P\bar{x})_P$$

The last equality follows from the fact that P is orthogonal with respect to the inner product being used. Thus

$$\|\bar{x}\|_P^2 = \|P\bar{x}\|_P^2 + \|\bar{x} - P\bar{x}\|_P^2 \geq \|P\bar{x}\|_P^2 .$$

Hence

$$\|P\bar{x}\|_P \leq \|\bar{x}\|_P$$

for all $\bar{x}$ in E_n. Thus $\|P\|_P \leq 1$. Therefore $\|P\|_P = 1$.

Lemma 2.1: Let P be an orthogonal projection with respect to a given inner product and suppose $\bar{x}$ is a vector in E_n such that $\|\bar{x}\| = \|P\bar{x}\|$ where $\|\cdot\|$ is the norm generated by the given inner product. Then $\bar{x}$ is in the range of P and $P\bar{x} = \bar{x}$.

Proof:

$$\|\bar{x}\|^2 = (\bar{x} - P\bar{x} + P\bar{x}, \bar{x} - P\bar{x} + P\bar{x})$$

$$\|\bar{x}\|^2 = (P\bar{x}, P\bar{x}) + (P\bar{x}, \bar{x} - P\bar{x}) + (\bar{x} - P\bar{x}, P\bar{x}) + (\bar{x} - P\bar{x}, \bar{x} - P\bar{x})$$

$$\|\bar{x}\|^2 = (P\bar{x}, P\bar{x}) + (\bar{x} - P\bar{x}, \bar{x} - P\bar{x}) = \|P\bar{x}\|^2 + \|\bar{x} - P\bar{x}\|^2$$

The last equality follows from the fact that P is an orthogonal projection with respect to the above inner product. From the hypothesis $\|\bar{x}\| = \|P\bar{x}\|$, so that $\|\bar{x} - P\bar{x}\| = 0$. Thus

$$\bar{x} = P\bar{x},$$

and $\bar{x}$ is in the range of P.

Theorem 2.4: Suppose $P_1, P_2, \ldots, P_n$ are orthogonal projections on E_n with respect to a given inner product on E_n. Let $R(P_i)$ denote the range of P_i, $1 \leq i \leq n$. Suppose further that

$$\bigcap_{i=1}^{n} R(P_i) = \{\bar{0}\},$$

Then

$$\| \prod_{i=1}^{n} P_i \| < 1$$

where $\|\cdot\|$ is the norm generated by the given inner product.

Proof: Let $\bar{y}$ be a non-zero vector in E_n. Consider $P_k\cdots P_1\bar{y}$, $1 \leq k \leq n$. For $k \geq 2$

$$\|P_k P_{k-1}\cdots P_1\bar{y}\| \leq \|P_k\| \, \|P_{k-1}\cdots P_1\bar{y}\| \, .$$

Also

$$\|P_1\bar{y}\| \leq \|P_1\| \, \|\bar{y}\| \, .$$

But each P_i is an orthogonal projection so $\|P_i\| = 1$, $1 \leq i \leq n$. Thus

$$\|P_1\bar{y}\| \leq \|\bar{y}\|$$

and for $2 \leq k \leq n$,

$$\|P_k P_{k-1}\cdots P_1\bar{y}\| \leq \|P_{k-1}\cdots P_1\bar{y}\| \, .$$

Now suppose equality holds in all of the above. Applying Lemma 2.1 yields, for $2 \leq k \leq n$,

$$P_k\cdots P_1\bar{y} = P_{k-1}\cdots P_1\bar{y}$$

where $P_{k-1}\cdots P_1\bar{y}$ plays the role of the $\bar{x}$ in the lemma, and

$$P_1\bar{y} = \bar{y}$$

where $\bar{y}$ plays the role of $\bar{x}$ in the lemma. Putting all these equalities together yields

$$P_i\bar{y} = \bar{y}$$

for all i, $1 \leq i \leq n$. Thus $\bar{y}$ is in $R(P_i)$ for all i and hence

$$\bar{y} \text{ is in } \bigcap_{i=1}^{n} R(P_i) \ .$$

But this holds only if $\bar{y} = \bar{0}$ which is impossible since $\bar{y}$ was chosen non-zero. Thus the assumption is false and it must be the case that

$$\|P_1\bar{y}\| < \|\bar{y}\|$$

or

$$\|P_k \cdots P_1\bar{y}\| < \|P_{k-1} \cdots P_1\bar{y}\|$$

for some k, $2 \leq k \leq n$. This yields

$$\| \prod_{i=1}^{n} P_i\bar{y}\| < \|\bar{y}\|$$

for all $\bar{y}$ in E_n. In particular, for $\bar{y}$ such that $\|\bar{y}\| = 1$,

$$\| \prod_{i=1}^{n} P_i\bar{y}\| < 1 \ .$$

Now the function $g(\bar{y}) = \| \prod_{i=1}^{n} P_i\bar{y}\|$ is a continuous function on E_n and the set $\left\{\bar{x} : \|\bar{x}\| = 1\right\}$ is a compact set in E_n. Thus there exists an $\bar{x}_0$ with norm one such that

$$\| \prod_{i=1}^{n} P_i \ \bar{x}_0\| = \text{Sup}\left\{\| \prod_{i=1}^{n} P_i\bar{y}\| : \|\bar{y}\| = 1\right\} .$$

Recalling the definition of $\| \prod_{i=1}^{n} P_i\|$, one then has

$$\left\| \prod_{i=1}^{n} P_i \right\| = \left\| \prod_{i=1}^{n} P_i \ \bar{x}_0 \right\| < \|\bar{x}_0\| = 1 \ .$$

Theorem 2.5: Suppose the P_t's in a projective iterative scheme (1.1) are such that there exists a positive integer m such that $P_{t+m} = P_t$ for all $t \geq 0$. Suppose also that for $0 \leq t \leq m-1$, P_t is an orthogonal projection with respect to some fixed inner product and that

$$\bigcap_{i=0}^{m-1} R(P_i) = \left\{ \bar{0} \right\},$$

where $R(P_i)$ denotes the range of P_i. Then the iterative scheme (1.1) will converge to the solution of $A\bar{x} = \bar{b}$.

Proof: Since for each t, P_t is an orthogonal projection with respect to a given inner product, then, using the norm generated by that inner product, one has

$$\|P_t\| = 1 \ , \quad t \geq 0 \ .$$

Now since $P_{t+m} = P_t$ for all $t \geq 0$ and $\bigcap_{i=0}^{m-1} R(P_i) = \left\{ \bar{0} \right\}$,

$$\bigcap_{i=0}^{m-1} R(P_{k+i}) = \left\{ \bar{0} \right\},$$

where k is an integer greater than or equal to zero. Also the P_i are orthogonal projections with respect to a given inner product. Thus Theorem 2.4 applies and one has

$$\left\| \prod_{i=0}^{m-1} P_{k+i} \right\| < 1$$

for $k \geq 0$. Define α as

$$\alpha = \mathrm{Max}\left\{\left\|\prod_{i=0}^{m-1} P_{k+i}\right\| : 0 \leq k \leq m-1\right\}.$$

Note that $\alpha < 1$. Now since $P_{t+m} = P_t$ one has

$$\left\|\prod_{i=0}^{m-1} P_{t+i}\right\| < \alpha < 1$$

for all $t \geq 0$. Thus Theorem 2.2 applies and the iterative scheme (1.1) converges.

The following two corollaries show how the last theorem can be used to test for convergence of a projective iterative scheme.

<u>Corollary 2.5.1</u>: If A is a positive definite Hermitian matrix, then the P_t associated with the Gauss-Seidel iterative scheme for solving the system $A\bar{x} = \bar{b}$ are A-orthogonal and the Gauss-Seidel iterative scheme converges.

Proof: Recall from Example I, Chapter I, that the P_t associated with the Gauss-Seidel scheme are given by $P_t = I - C_tA$ where for $0 \leq t \leq n-1$,

$$C_t = \left[\begin{array}{c|c|c} 0_t & 0 & 0 \\ \hline 0 & \dfrac{1}{a_{t+1t+1}} & 0 \\ \hline 0 & 0 & 0_{n-t-1} \end{array}\right]$$

and for $t \geq n$,

$$C_t = C_{t \bmod n}.$$

Now let t be such that $1 \leq t \leq n$.

It can be easily seen (recalling the expression for $C_{t-1}A$ in Example I, Chapter I) that for a vector $\overline{x}$

$$(2.1) \qquad P_{t-1}\overline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{t-1} \\ -\sum\limits_{\substack{j=1 \\ j\neq t}}^{n} \frac{a_{tj}}{a_{tt}} x_j \\ x_{t+1} \\ \vdots \\ x_n \end{bmatrix} .$$

Thus the range of P_{t-1}, $R(P_{t-1})$, consists of vectors of the form on the right above.

It will now be shown that the above P_{t-1}'s, $1 \leq t \leq n$, are A-orthogonal. This will be true if

$$(\overline{x} - P_{t-1}\overline{x})^* AP_{t-1}\overline{x} = 0 .$$

It is easily seen using (2.1) that

$$\overline{x} - P_{t-1}\overline{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_t + \sum\limits_{\substack{j=1 \\ j\neq t}}^{n} \frac{a_{tj}}{a_{tt}} x_j \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

Again using (2.1) it is easily seen that

$$AP_{t-1}\overline{x} = \begin{bmatrix} \sum\limits_{\substack{k=1\\k\neq t}}^{n} a_{1k}x_k - \sum\limits_{\substack{j=1\\j\neq t}}^{n} \frac{a_{1t}a_{tj}}{a_{tt}} x_j \\ \vdots \\ \sum\limits_{\substack{k=1\\k\neq t}}^{n} a_{tk}x_k - \sum\limits_{\substack{j=1\\j\neq t}}^{n} \frac{a_{tt}a_{tj}}{a_{tt}} x_j \\ \vdots \\ \sum\limits_{\substack{k=1\\k\neq t}}^{n} a_{nk}x_k - \sum\limits_{\substack{j=1\\j\neq t}}^{n} \frac{a_{nt}a_{tj}}{a_{tt}} x_j \end{bmatrix} .$$

Note that the t^{th} element of $AP_{t-1}\overline{x}$ turns out to be zero. But all of the elements of $\overline{x} - P_{t-1}\overline{x}$ are zero except the t^{th}. Thus

$$(\overline{x} - P_{t-1}\overline{x})^* AP_{t-1}\overline{x} = 0 .$$

Hence P_{t-1} is A-orthogonal for $1 \leq t \leq n$. For $t \geq n$, $P_t = P_j$ for some $j \leq n - 1$, so P_t is A-orthogonal for all $t \geq 0$.

It will now be shown that $\bigcap\limits_{i=1}^{n} R(P_{i-1}) = \{\overline{0}\}$. Note that the i^{th} component of a vector in $R(P_{i-1})$ is of the form

$$-\sum_{\substack{j=1\\j\neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j .$$

Now suppose that $\overline{x}$ is in $\bigcap\limits_{i=1}^{n} R(P_{i-1})$. Then $\overline{x}$ is in $R(P_{i-1})$ for $1 \leq i \leq n$, and hence

$$x_i = - \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j, \quad 1 \leq i \leq n .$$

Then for $1 \leq i \leq n$,

$$x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{a_{ij}}{a_{ii}} x_j = 0$$

or

$$a_{ii}x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j = \sum_{j=1}^{n} a_{ij}x_j = 0 .$$

But these are just the equations that make up the system

$$A\bar{x} = \bar{0} .$$

Therefore, since A is non-singular, $\bar{x} = \bar{0}$. Thus one has that the P_t are orthogonal projections with respect to the same inner product, the one generated by A, and that

$$\bigcap_{i=0}^{n-1} R(P_i) = \{\bar{0}\}.$$

Also $P_{t+n} = P_t$ for all $t \geq 0$. Thus by Theorem 2.5 the iterative scheme (1.1), which in this case turns out to be the Gauss-Seidel iterative scheme, converges to the solution of $A\bar{x} = \bar{b}$.

Corollary 2.5.2: If A is a Hermitian positive definite matrix, then the block Gauss-Seidel iterative scheme associated with the system $A\bar{x} = \bar{b}$ converges.

Proof: For this scheme the matrix A is partitioned in the form

$$\begin{bmatrix} A_{11} \cdots A_{1N} \\ \vdots \quad\quad \vdots \\ A_{N1} \cdots A_{NN} \end{bmatrix},$$

where A_{ii} is square and non-singular for $1 \le i \le N$. Recall from Example II, Chapter I, that the projections P_t for this method are given by

$$P_t = I - C_t A ,$$

where for $0 \le t \le N-1$,

$$C_t = \left[\begin{array}{c|c|c} 0_{p_t} & 0 & 0 \\ \hline 0 & A^{-1}_{t+1\,t+1} & 0 \\ \hline 0 & 0 & 0_{n-p_{t+1}} \end{array}\right],$$

($p_t = \sum_{j=1}^{i} v_j$, where v_j is the number of rows in A_{jj}) and for $t \ge N$,

$$C_t = C_{t \bmod N} .$$

It will now be shown that P_{t-1} is A-orthogonal for all t, $1 \le t \le N$. Let $\bar{x}$ be a vector in E_n. This vector is partitioned in the form

$$\overline{x} = \begin{bmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_i \\ \vdots \\ \overline{x}_N \end{bmatrix}$$

where each $\overline{x}_i$ has the same number of components as A_{ii} has rows or columns. It can easily be seen recalling the expression for $C_{t-1}A$ in Example II, Chapter I, that

$$(2.2) \qquad P_{t-1}\overline{x} = \begin{bmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_{t-1} \\ -\sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{tt}^{-1}A_{tj}\overline{x}_j \\ \overline{x}_{t+1} \\ \vdots \\ \overline{x}_N \end{bmatrix}$$

In order to show that P_{t-1} is A-orthogonal, it must be shown that

$$(\overline{x} - P_{t-1}\overline{x})^* A P_{t-1}\overline{x} = 0 .$$

It is easily seen using (2.2) that

$$\bar{x} - P_{t-1}\bar{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \bar{x}_t + \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{tt}^{-1}A_{tj}\bar{x}_j \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Again using (2.2) it is easily seen that

$$AP_{t-1}\bar{x} = \begin{bmatrix} \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{1j}\bar{x}_j - \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{1t}A_{tt}^{-1}A_{tj}\bar{x}_j \\ \vdots \\ \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{tj}\bar{x}_j - \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{tt}A_{tt}^{-1}A_{tj}\bar{x}_j \\ \vdots \\ \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{Nj}\bar{x}_j - \sum\limits_{\substack{j=1 \\ j\neq t}}^{N} A_{Nt}A_{tt}^{-1}A_{tj}\bar{x}_j \end{bmatrix} .$$

Note that the t^{th} block of $AP_{t-1}\bar{x}$ turns out to be zero. But all the blocks of $\bar{x} - P_{t-1}\bar{x}$ are zero except the t^{th}. Thus

$$(\bar{x} - P_{t-1}\bar{x})^{*}AP_{t-1}\bar{x} = 0 .$$

Hence P_{t-1} is A-orthogonal for $1 \leq t \leq N$. Now for $t \geq N$ $P_t = P_j$ for some j, $0 \leq j \leq N - 1$, so P_t is A-orthogonal for all $t \geq 0$.

It will now be shown that $\bigcap_{i=1}^{N} R(P_{i-1}) = \{\overline{0}\}$. Here $R(P_{i-1})$ denotes the range of P_{i-1}. From (2.2) one has that the i^{th} block of the N-blocked partition of a vector $\overline{x}$ in the range of P_{i-1} is of the form

$$-\sum_{\substack{j=1 \\ j\neq i}}^{N} A_{ii}^{-1}A_{ij}\overline{x}_j ,$$

where the $\overline{x}_j$ are the blocks of the N-blocked partition of $\overline{x}$. Now suppose $\overline{x}$ is in $\bigcap_{i=1}^{N} R(P_{i-1})$. Then from the above

$$\overline{x}_i = -\sum_{\substack{j=1 \\ j\neq i}}^{N} A_{ii}^{-1}A_{ij}\overline{x}_j , \quad 1 \leq i \leq N .$$

Then for $1 \leq i \leq N$,

$$\overline{x}_i + \sum_{\substack{j=1 \\ j\neq i}}^{N} A_{ii}^{-1}A_{ij}\overline{x}_j = \overline{0} ,$$

or

$$A_{ii}\overline{x}_i + \sum_{\substack{j=1 \\ j\neq i}}^{N} A_{ij}\overline{x}_j = \sum_{j=1}^{N} A_{ij}\overline{x}_j = \overline{0} .$$

But these are just the blocks of equations that make up the system

$$A\overline{x} = \overline{0} .$$

Therefore, since A is non-singular, $\bar{x} = \bar{0}$.

Thus one sees that the P_t are orthogonal projections with respect to the same inner product, the one generated by A, and that

$$\bigcap_{i=0}^{N-1} R(P_i) = \{\bar{0}\}.$$

Also $P_{t+N} = P_t$ for all $t \geq 0$. Thus by Theorem 2.5 the iterative scheme (1.1), which in this case is the block Gauss-Seidel iterative scheme, converges to the solution of $A\bar{x} = \bar{b}$.

CHAPTER III

NORM REDUCING METHODS

In the general setting it is difficult to get anything more than vague or general results concerning the behavior of projective iterative schemes. However, by restricting discussion to a particular class of projective iterative schemes, considerable knowledge may be gained about the behavior of these special methods. This concluding chapter deals with one such class of projective iterative schemes, those which are norm reducing schemes. The behavior of such schemes can be studied in greater detail than that of projective schemes in general.

Definition 3.1: A norm reducing projective iterative scheme is a projective iterative method of the form (1.1) for solving the system $A\bar{x} = \bar{b}$ such that the error vectors at each iteration, denoted by $\bar{s}_{t+1} = \bar{u} - \bar{x}_{t+1}$ where $\bar{u}$ is the solution of the system, satisfy the relation

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \;, \quad t \geq 0 \;,$$

for some fixed norm.

The method of steepest descent and the conjugate directions iterative schemes, Examples III and IV in Chapter I, give examples of norm reducing iterative schemes.

General Discussion

The discussion in this chapter deals with elliptic norms, i.e.,

$$\|\bar{y}\|_G = (\bar{y}^* G \bar{y})^{1/2}$$

where G is a positive definite Hermitian matrix. In addition to this restriction the discussion will further be restricted to deal only with those norm reducing schemes of the form of the general method given in the following theorem.

Theorem 3.1: Let $\bar{u}$ be the solution of $A\bar{x} = \bar{b}$ and Y_t be an $n \times k_t$ $(1 \leq k_t \leq n)$ matrix with k_t linearly independent columns. Consider the scheme

$$\bar{x}_{t+1} = \bar{x}_t + Y_t \bar{v}_t = \bar{x}_t + C_t(\bar{b} - A\bar{x}_t) \tag{3.1}$$

where (i) $\bar{v}_t = (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t$, where G is a positive definite Hermitian matrix and $\bar{s}_t = \bar{u} - \bar{x}_t$

and (ii) C_t is a matrix which satisfies

$$C_t(\bar{b} - A\bar{x}_t) = Y_t \bar{v}_t .$$

Then the scheme (3.1) is a norm reducing (with respect to the norm generated by G) projective iterative scheme;

$$\|\bar{s}_t\|_G^2 - \|\bar{s}_{t+1}\|_G^2 = \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t \;;$$

and

$$C_t = Y_t (Y_t^* G Y_t)^{-1} Y_t^* G A^{-1}$$

for all $t \geq 0$.

Proof: Define the function $f(\bar{y})$ for vectors $\bar{y}$ by

$$f(\bar{y}) = \|\bar{s}_t - Y_t\bar{y}\|_G^2 = (\bar{s}_t - Y_t\bar{y})^* G(\bar{s}_t - Y_t\bar{y}) \text{ .}$$

Note $f(\bar{0}) = \|\bar{s}_t\|_G^2$. Now one minimizes $f(\bar{y})$. In order to do this one lets

$$\bar{y} = (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t + \bar{w}_t$$

($(Y_t^* G Y_t)^{-1}$ exists by Theorem 1.3). The following computations result from substituting the above expression for $\bar{y}$ in $f(\bar{y})$ and simplifying, using the fact that

$$((Y_t^* G Y_t)^{-1})^* = ((Y_t^* G Y_t)^*)^{-1} = (Y_t^* G Y_t)^{-1}$$

as $G = G^*$ and $Y_t^{**} = Y_t$.

$$f(\bar{y}) = \bar{s}_t^* G \bar{s}_t - \bar{y}^* Y_t^* G \bar{s}_t - \bar{s}_t^* G Y_t \bar{y} + \bar{y}^* Y_t^* G Y_t \bar{y}$$

$$\begin{aligned} f(\bar{y}) = {} & \bar{s}_t^* G \bar{s}_t - \bar{s}_t^* G^* Y_t ((Y_t^* G Y_t)^{-1})^* Y_t G \bar{s}_t - \bar{w}_t^* Y_t^* G \bar{s}_t \\ & - \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t - \bar{s}_t^* G Y_t^* \bar{w}_t + \bar{w}_t^* Y_t^* G Y_t \bar{w}_t \\ & + \bar{s}_t^* G^* Y_t ((Y_t^* G Y_t)^{-1})^* Y_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t \\ & + \bar{s}_t^* G^* Y_t ((Y_t^* G Y_t)^{-1})^* Y_t^* G Y_t \bar{w}_t + \bar{w}_t^* Y_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t \end{aligned}$$

$$\begin{aligned} f(\bar{y}) = {} & \bar{s}_t^* G \bar{s}_t - \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t - \bar{w}_t^* Y_t^* G \bar{s}_t + \bar{w}_t^* Y_t^* G Y_t \bar{w}_t \\ & - \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t - \bar{s}_t^* G Y_t \bar{w}_t + \bar{s}_t^* G Y_t (Y^* G Y_t)^{-1} Y_t^* G \bar{s}_t \\ & + \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G Y_t \bar{w}_t + \bar{w}_t^* Y_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t \end{aligned}$$

$$f(\bar{y}) = \bar{s}_t^* G\bar{s}_t - \bar{w}_t^* Y_t^* G\bar{s}_t + \bar{w}_t^* Y_t^* GY_t\bar{w}_t - \bar{s}_t^* GY_t(Y_t^* GY_t)^{-1}Y_t^* G\bar{s}_t$$

$$- \bar{s}_t^* GY_t\bar{w}_t + \bar{s}_t^* GY_t\bar{w}_t + \bar{w}_t^* Y_t^* G\bar{s}_t$$

$$(3.2) \qquad f(\bar{y}) = \bar{s}_t^* G\bar{s}_t - \bar{s}_t^* GY_t(Y_t^* GY_t)^{-1}Y_t^* G\bar{s}_t + \bar{w}_t^* Y_t^* GY_t\bar{w}_t$$

Now $Y_t^* GY_t$ is a positive definite matrix, by Theorem 1.3, so $\bar{w}_t^* Y_t^* GY_t\bar{w}_t \geq 0$ with equality only when $\bar{w} = 0$. But note that $f(\bar{y})$ depends only on $\bar{w}_t$ as everything else is fixed. Thus $f(\bar{y})$ is minimized by choosing $\bar{w} = \bar{0}$, i.e., for

$$\bar{y} = (Y_t^* GY_t)^{-1}Y_t^* G\bar{s}_t \ .$$

But this is the $\bar{v}_t$ defined in the hypothesis of the theorem. Therefore,

$$f(\bar{v}_t) \leq f(\bar{y})$$

for all $\bar{y}$ including $\bar{y} = \bar{0}$. But $f(\bar{0}) = \|\bar{s}_t\|_G^2$. From (3.1)

$$\bar{s}_t - Y_t\bar{v}_t = \bar{u} - (\bar{x}_t + Y_t\bar{v}_t) = \bar{u} - \bar{x}_{t+1} = \bar{s}_{t+1} \ .$$

Thus,

$$\|\bar{s}_{t+1}\|_G^2 = \bar{s}_{t+1}^* G\bar{s}_{t+1} = (\bar{s}_t - Y_t\bar{v}_t)^* G(\bar{s}_t - Y_t\bar{v}_t) = f(\bar{v}_t)$$

and

$$\|\bar{s}_{t+1}\|_G^2 = f(\bar{v}_t) \leq f(\bar{0}) = \|\bar{s}_t\|_G^2 \ .$$

In fact, from (3.2), since $\bar{w} = \bar{0}$ for $\bar{y} = \bar{v}_t$,

$$\|\bar{s}_{t+1}\|_G^2 = f(\bar{v}_t) = \|\bar{s}_t\|_G^2 - \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t ,$$

or

$$\|\bar{s}_t\|_G^2 - \|\bar{s}_{t+1}\|_G^2 = \bar{s}_t^* G Y_t (Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t .$$

Now $C_t(\bar{b} - A\bar{x}_t) = Y_t\bar{v}_t$ by hypothesis and $\bar{b} = A\bar{u}$. Thus

$$Y_t\bar{v}_t = C_t(A\bar{u} - A\bar{x}_t) = C_t A(\bar{u} - \bar{x}_t) = C_t A\bar{s}_t .$$

Hence,

$$C_t A\bar{s}_t = Y_t(Y_t^* G Y_t)^{-1} Y_t^* G \bar{s}_t .$$

But this must hold for all $\bar{s}_t$ so that

$$C_t A = Y_t(Y_t^* G Y_t)^{-1} Y_t^* G .$$

Therefore,

$$C_t = Y_t(Y_t^* G Y_t)^{-1} Y_t^* G A^{-1} .$$

Now from (3.1),

$$\bar{x}_{t+1} = \bar{x}_t + C_t(\bar{b} - A\bar{x}_t) = (I - C_t A)\bar{x}_t + C_t\bar{b} .$$

Thus (3.1) is a projective iterative scheme if $I - C_t A$ is a projection. Note that

$$I - C_t A = I - Y_t(Y_t^* G Y_t)^{-1} Y_t^* G .$$

Thus by Theorem 1.4, $I - C_t A$ is a projection, and hence (3.1) defines a norm reducing projective iterative scheme. This completes the proof.

Note that the above method is feasible only if the C_t's are determinable without computing A^{-1}, i.e., a matrix V_t can be easily found to satisfy the relation

$$Y_t^* G = V_t^* A .$$

There are, however, several ways to make the scheme (3.1) usable, and the remainder of this chapter explores these methods and the behavior of the resulting iterative schemes. Basically all the methods used employ a choice of G which makes it easy to compute V_t in terms of Y_t or else using a G where G^{-1} is known, choosing the V_t, and then computing the Y_t.

The following corollaries give schemes resulting from various choices of G and restrictions on Y_t. Later the choice $G = A$ when A is a positive definite Hermitian matrix will be considered.

<u>Corollary 3.1.1</u>: The iterative scheme

$$(3.3) \qquad \bar{x}_{t+1} = (I - Y_t(Y_t^* A^* A Y_t)^{-1} Y_t^* A^*)\bar{x}_t + Y_t(Y_t^* A^* A Y_t)^{-1} Y_t^* A^* \bar{b}$$

is a norm reducing projective iterative scheme and

$$\|\bar{s}_t\|^2_{A^*A} - \|\bar{s}_{t+1}\|^2_{A^*A} = \bar{r}_t^* \, A Y_t (Y_t^* A^* A Y_t)^{-1} Y_t^* A^* \bar{r}_t$$

where $\bar{r}_t = \bar{b} - A\bar{x}_t$ is the residual vector.

Proof: First note that $\bar{r}_t = A\bar{s}_t$ so that

$$\bar{r}_t^* A Y_t (Y_t^* A^* A Y_t)^{-1} Y_t^* A^* \bar{r}_t = \bar{s}_t^* A^* A Y_t (Y_t^* A^* A Y_t)^{-1} Y_t^* A^* A \bar{s}_t .$$

Using this fact, the conclusion of this corollary is just that of Theorem 3.1 with $G = A^*A$.

Corollary 3.1.2: The iterative scheme

$$(3.4) \qquad \bar{x}_{t+1} = (I - A^*V_t(V_t^*AA^*V_t)^{-1}V_t^*)\bar{x}_t + A^*V_t(V_t^*AA^*V_t)^{-1}V_t^*\bar{b}$$

is a norm reducing projective iterative scheme and

$$\|\bar{s}_t\|^2 - \|\bar{s}_{t+1}\|^2 = \bar{r}_t^*V_t(V_t^*AA^*V_t)^{-1}V_t^*\bar{r}_t .$$

Proof: First, since $\bar{r}_t = A\bar{s}_t$, note that

$$\bar{r}_t^*V_t(V_t^*AA^*V_t)^{-1}V_t^*\bar{r}_t = \bar{s}_t^*A^*V_t(V_t^*AA^*V_t)^{-1}V_t^*A\bar{s}_t .$$

Using this fact, the conclusion of this corollary is just that of Theorem 3.1 with $G = I$, and $Y_t = A^*V_t$. (Note that with this scheme the V_t are chosen and the Y_t are determined from them.)

Corollary 3.1.3: The iterative scheme

$$(3.5) \qquad \bar{x}_{t+1} = \left[I - \frac{\bar{y}_t\bar{y}_t^*G}{\bar{y}_t^*G\bar{y}_t} \right] \bar{x}_t + \frac{\bar{y}_t\bar{y}_t^*GA^{-1}\bar{b}}{\bar{y}_t^*G\bar{y}_t}$$

is a norm reducing projective iterative scheme and

$$\|\bar{s}_t\|_G^2 - \|\bar{s}_{t+1}\|_G^2 = \frac{\bar{s}_t^*G\bar{y}_t\bar{y}_t^*G\bar{s}_t}{\bar{y}_t^*G\bar{y}_t} ,$$

Proof: The conclusion of this corollary is just that of Theorem 3.1 with $Y_t = \bar{y}_t$.

Corollary 3.1.4: The iterative scheme

$$(3.6) \qquad \bar{x}_{t+1} = \left[I - \frac{\bar{y}_t \bar{y}_t^* A^* A}{\bar{y}_t^* A^* A \bar{y}_t} \right] \bar{x}_t + \frac{\bar{y}_t \bar{y}_t^* A^* \bar{b}}{\bar{y}_t^* A^* A \bar{y}_t}$$

is a norm reducing projective iterative scheme and

$$\|\bar{s}_t\|^2_{A^*A} - \|\bar{s}_{t+1}\|^2_{A^*A} = \frac{\bar{r}_t^* A \bar{y}_t \bar{y}_t^* A^* \bar{r}_t}{\bar{y}_t^* A^* A \bar{y}_t},$$

where $\bar{r}_t = \bar{b} - A\bar{x}_t = A\bar{s}_t$.

Proof: Using the fact that $\bar{r}_t = A\bar{s}_t$, one has

$$\frac{\bar{r}_t^* A \bar{y}_t \bar{y}_t^* A^* \bar{r}_t}{\bar{y}_t^* A^* A \bar{y}_t} = \frac{\bar{s}_t^* A^* A \bar{y}_t \bar{y}_t^* A^* A \bar{s}_t}{\bar{y}_t^* A^* A \bar{y}_t}.$$

This equation shows that the conclusion of the corollary is just that of Corollary 3.1.3 with $G = A^*A$.

Corollary 3.1.5: The iterative scheme

$$(3.7) \qquad \bar{x}_{t+1} = \left[I - \frac{A^* \bar{v}_t \bar{v}_t^* A}{\bar{v}_t^* A A^* \bar{v}_t} \right] \bar{x}_t + \frac{A^* \bar{v}_t \bar{v}_t^* \bar{b}}{\bar{v}_t^* A A^* \bar{v}_t}$$

is a norm reducing projective iterative scheme and

$$\|\bar{s}_t\|^2 - \|\bar{s}_{t+1}\|^2 = \frac{\bar{r}_t^* \bar{v}_t \bar{v}_t^* \bar{r}_t}{\bar{v}_t^* A A^* \bar{v}_t},$$

where $\bar{r}_t = A\bar{s}_t = \bar{b} - A\bar{x}_t$.

Proof: The conclusion of this corollary is just that of Corollary 3.1.2 with $V_t = \bar{v}_t$.

Behavior of Schemes

Before proceeding further with the behavior of certain norm reducing projective iterative schemes, it is first necessary to state some inequalities which will be used in the succeeding theorems.

Theorem 3.2 (Inequality of Wielandt): Let $k(A) = \|A\| \, \|A^{-1}\|$ be the condition number of the matrix A and $\bar{x}$ and $\bar{y}$ be any pair of orthonormal vectors, i.e., $\bar{x}^*\bar{y} = 0$. Define θ to be an angle which satisfies

$$\cos\theta = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Then the following relation holds.

$$\frac{|\bar{x}^*A^*A\bar{y}|}{\bar{x}^*A^*A\bar{x}\bar{y}^*A^*A\bar{y}} \leq \cos^2\theta .$$

A proof of this can be found in Householder [6].

Theorem 3.3 (Generalized Wielandt Inequality): If the vectors $\bar{x}$ and $\bar{y}$ enclose an angle of not less than φ, then for a non-singular matrix A, $A\bar{x}$ and $A\bar{y}$ enclose an angle φ_A satisfying

$$\cot\frac{\varphi_A}{2} \leq k(A)\,\cot\frac{\varphi}{2} ,$$

where all the angles are taken in the first quadrant and $k(A)$ denotes the condition number of A. The angle φ satisfies the relation

$$\cos \varphi = \frac{\bar{x}^* \bar{y}}{(\bar{x}^* \bar{x})(\bar{y}^* \bar{y})} .$$

Defining the angle Ψ by the relation

$$\cot \frac{\Psi}{2} = k(A) \cot \frac{\varphi}{2}$$

gives the inequality

$$\frac{|\bar{x}^* A^* A \bar{y}|^2}{\bar{x}^* A^* A \bar{x} \bar{y}^* A^* A \bar{y}} \leq \cos^2 \Psi .$$

A proof of this can be found in Householder [7].

<u>Theorem 3.4</u> <u>(The Kantorovich Inequality)</u>: Let A be a non-singular matrix and $k(A)$ be the condition number of A. Let $\bar{x}$ be a vector and define θ to satisfy

$$\cos \theta = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Then

$$\frac{(\bar{x}^* \bar{x})^2}{\bar{x}^* A^* A \bar{x} \bar{x}^* (A^* A)^{-1} \bar{x}} \geq \sin^2 \theta .$$

A proof of this can be found in Householder [7].

Now the behavior of various norm reducing projective iterative schemes will be studied.

<u>Theorem 3.5</u>: Consider the iterative scheme (3.4) associated with the system $A\bar{x} = \bar{b}$ and recall that the V_t are $n \times k_t$ $(1 \leq k_t \leq n)$ matrices

such that the columns are linearly independent. Let $k(A)$ be the condition number of the matrix A. Let φ be the first quadrant angle which satisfies

$$\sin \varphi = \frac{\|V_t(V_t^*V_t)^{-1}V_t^*\bar{r}_t\|}{\|\bar{r}_t\|} .$$

Then (i) the iteration (3.4) reduces the norm of the error vector

$\bar{s}_t = \bar{u} - \bar{x}_t$, where $\bar{u}$ is the solution of $A\bar{x} = \bar{b}$, and, in fact,

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \quad \cos \Psi ,$$

where $\cot \frac{\Psi}{2} = k(A) \cot \frac{\varphi}{2}$;

and (ii) if, further, $\bar{r}_t$ is contained in the space generated by the columns of V_t, then

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Proof: Let $P_t = I - C_tA$, where from (3.4)

$$C_t = A^*V_t(V_t^*AA^*V_t)^{-1}V_t^* .$$

It will now be shown that P_t is a self-adjoint projection. It is already known that P_t is a projection so it need only be shown that it is self-adjoint.

$$P_t^* = (I - A^*V_t(V_t^*AA^*V_t)^{-1}V_t^*A)^*$$

$$P_t^* = I^* - A^*(V_t^*)^*((V_t^* AA^* V_t)^{-1})^* V_t^*(A^*)^*$$

$$P_t^* = I - A^* V_t((V_t^* AA^* V_t)^*)^{-1} V_t^* A$$

$$P_t^* = I - A^* V_t(V_t^*(A^*)^* A^*(V_t^*)^*)^{-1} V_t^* A$$

$$P_t^* = I - A^* V_t(V_t^* AA^* V_t)^{-1} V_t^* A = P_t \ .$$

Thus P_t is self-adjoint.

Let $\bar{w}_t = AP_t\bar{s}_t$. It will be shown that

$$\|\bar{s}_{t+1}\|^2 = \bar{r}_t^*(AA^*)^{-1}\bar{w}_t \ .$$

$$\|\bar{s}_{t+1}\|^2 = \|\bar{u} - \bar{x}_{t+1}\|^2$$

$$\|\bar{s}_{t+1}\|^2 = \|\bar{u} - (I - C_t A)\bar{x}_t - C_t\bar{b}\|^2$$

Note that $\bar{b} = A\bar{u}$.

$$\|\bar{s}_{t+1}\|^2 = \|\bar{u} - \bar{x}_t - C_t(A\bar{u} - A\bar{x}_t)\|^2$$

$$\|\bar{s}_{t+1}\|^2 = \|\bar{s}_t - C_t A\bar{s}_t\|^2 = \|(I - C_t A)\bar{s}_t\|^2$$

$$\|\bar{s}_{t+1}\|^2 = \|P_t\bar{s}_t\|^2 = \bar{s}_t^* P_t^* P_t \bar{s}_t = \bar{s}_t^* P_t P_t \bar{s}_t$$

$$\|\bar{s}_{t+1}\|^2 = \bar{s}_t^* P_t \bar{s}_t = \bar{s}_t^* A^*(A^*)^{-1}A^{-1}AP_t\bar{s}_t$$

Note that $A\bar{s}_t = \bar{r}_t$ and $\bar{r}_t^* = \bar{s}_t^* A^*$. Therefore

$$\|\bar{s}_{t+1}\|^2 = \bar{r}_t^*(AA^*)^{-1}AP_t\bar{s}_t = \bar{r}_t^*(AA^*)^{-1}\bar{w}_t \ .$$

It will now be shown that $\bar{w}_t$ is orthogonal to the columns of V_t. This will be true if $\bar{w}_t^* V_t = \bar{0}$.

$$\bar{w}_t^* V_t = \bar{s}_t^* P_t^* A^* V_t = \bar{s}_t^* P_t A^* V_t$$

$$\bar{w}_t^* V_t = \bar{s}_t^*(I - A^*V_t(V_t^*AA^*V_t)^{-1}V_t^*A)A^*V_t$$

$$\bar{w}_t^* V_t = \bar{s}_t^* A^* V_t - \bar{s}_t^* A^* V_t (V_t^* AA^* V_t)^{-1} V_t^* AA^* V_t$$

$$\bar{w}_t^* V_t = \bar{s}_t^* A^* V_t - \bar{s}_t^* A^* V_t I = \bar{0}$$

Thus $\bar{w}_t$ is orthogonal to the columns of V_t.

It will now be shown that

$$\bar{r}_t^*(AA^*)^{-1}\bar{w}_t = \bar{w}_t^*(AA^*)^{-1}\bar{w}_t \ .$$

$$\bar{0} = \bar{s}_t^* P_t \bar{s}_t - \bar{s}_t^* P_t P_t \bar{s}_t = (\bar{s}_t^* - \bar{s}_t^* P_t) P_t \bar{s}_t$$

since $P_t = P_t P_t$. Note that $P_t = P_t^*$ and $\bar{w}_t = AP_t\bar{s}_t$.

$$\bar{0} = (\bar{s}_t^* - \bar{s}_t^* P_t^*) A^* (A^*)^{-1} A^{-1} A P_t \bar{s}_t = (\bar{s}_t^* A^* - \bar{s}_t^* P_t^* A^*)(AA^*)^{-1}\bar{w}_t$$

But $\bar{r}_t^* = \bar{s}_t^* A^*$ and $\bar{w}_t^* = \bar{s}_t^* P_t^* A^*$. Therefore

$$\bar{0} = (\bar{r}_t^* - \bar{w}_t^*)(AA^*)^{-1}\bar{w}_t = \bar{r}_t^*(AA^*)^{-1}\bar{w}_t - \bar{w}_t^*(AA^*)^{-1}\bar{w}_t \ ,$$

or

$$\bar{r}_t^*(AA^*)^{-1}\bar{w}_t = \bar{w}_t^*(AA^*)^{-1}\bar{w}_t \ .$$

Now an expression will be developed for the ratio

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} \ .$$

Note that $\bar{r}_t = A\bar{s}_t$ so that $\bar{s}_t = A^{-1}\bar{r}_t$.

$$\|\bar{s}_t\|^2 = \bar{s}_t^*\bar{s}_t = \bar{r}_t^*(A^{-1})^*A^{-1}\bar{r}_t = \bar{r}_t^*(AA^*)^{-1}\bar{r}_t$$

It was shown previously that

$$\|\bar{s}_{t+1}\|^2 = \bar{r}_t^*(AA^*)^{-1}\bar{w}_t \ .$$

Thus,

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} = \frac{\bar{r}_t^*(AA^*)^{-1}\bar{w}_t}{\bar{r}_t^*(AA^*)^{-1}\bar{r}_t} = \frac{\bar{r}_t^*(AA^*)^{-1}\bar{w}_t \ \bar{w}_t^*(AA^*)^{-1}\bar{w}_t}{\bar{r}_t^*(AA^*)^{-1}\bar{r}_t \ \bar{w}_t^*(AA^*)^{-1}\bar{w}_t} \ .$$

But $\bar{w}_t^*(AA^*)^{-1}\bar{w}_t = \bar{r}_t^*(AA^*)^{-1}\bar{w}_t$. Note that this implies that both quantities are real since the first one is. Hence,

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} = \frac{(\bar{r}_t^*(AA^*)^{-1}\bar{w}_t)^2}{\bar{r}_t^*(AA^*)^{-1}\bar{r}_t \ \bar{w}_t^*(AA^*)^{-1}\bar{w}_t}$$

or

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} = \frac{|\bar{r}_t^*(AA^*)^{-1}\bar{w}_t|^2}{\bar{r}_t^*(AA^*)^{-1}\bar{r}_t\bar{w}_t^*(AA^*)^{-1}\bar{w}_t} \ .$$

Applying the generalized Wielandt inequality, Theorem 3.3, if the angle between $\bar{r}_t$ and $\bar{w}_t$ is at least φ', then

$$\frac{\|\bar{s}_{t+1}\|}{\|\bar{s}_t\|} \leq \cos \Psi ,$$

where Ψ is a first quadrant angle such that

$$\cot \frac{\Psi}{2} = k(A) \cot \frac{\varphi'}{2} .$$

Now $\bar{w}_t$ is orthogonal to the space generated by the columns of V_t and hence makes an angle of $\frac{\pi}{2}$ with it. Let φ, $0 \leq \varphi \leq \frac{\pi}{2}$, be such that

$$\sin \varphi = \frac{\|V_t(V_t^* V_t)^{-1} V_t^* \bar{r}_t\|}{\|\bar{r}_t\|} .$$

Then $\frac{\pi}{2} - \varphi$ is the angle between $\bar{r}_t$ and the space of V_t. Thus since $\bar{w}_t$ makes an angle of $\frac{\pi}{2}$ with the space of V_t and $\bar{r}_t$ makes an angle of $\frac{\pi}{2} - \varphi$ with that space, the angle between $\bar{r}_t$ and $\bar{w}_t$ cannot be less than $\frac{\pi}{2} - (\frac{\pi}{2} - \varphi) = \varphi$. Hence, taking $\varphi' = \varphi$,

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \cos \Psi ,$$

where $\cot \frac{\Psi}{2} = k(A) \cot \frac{\varphi}{2}$. This proves (i).

Now suppose $\bar{r}_t$ is contained in the space generated by V_t. Then the angle between $\bar{r}_t$ and the space of V_t is zero. Hence, $\varphi = \frac{\pi}{2}$. Therefore,

$$\cot \frac{\Psi}{2} = k(A) \cot \frac{\pi}{4} = k(A).$$

Now

$$\cos \Psi = \cos^2 \frac{\Psi}{2} - \sin^2 \frac{\Psi}{2} = \sin^2 \frac{\Psi}{2} \left(\cot^2 \frac{\Psi}{2} - 1\right) .$$

But $\cot^2 \frac{\Psi}{2} = (k(A))^2$ and

$$\sin^2 \frac{\Psi}{2} = \frac{1}{\csc^2 \frac{\Psi}{2}} = \frac{1}{1 + \cot^2 \frac{\Psi}{2}} = \frac{1}{1 + (k(A))^2} .$$

Thus,

$$\cos \Psi = \frac{(k(A))^2 - 1}{(k(A))^2 + 1} = \frac{k(A) \left[k(A) - \frac{1}{k(A)}\right]}{k(A) \left[k(A) + \frac{1}{k(A)}\right]}$$

or

$$\cos \Psi = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Therefore,

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

This proves (ii).

Corollary 3.5.1: Consider the scheme (3.7) associated with the system $A\bar{x} = \bar{b}$ where the $\bar{v}_t$ are chosen so that at the t^{th} stage $\bar{v}_t$ is that natural unit vector, $\bar{e}_i$, which maximizes the quantity $|\bar{r}_t^* \bar{e}_i|$. Then

$$\|\bar{s}_{t+1}\|^2 \leq \|\bar{s}_t\|^2 \left[1 - \frac{1}{n(k(A))^2}\right].$$

Proof: Since $\bar{e}_i$ is chosen so that $|\bar{r}_t^* \bar{e}_i|$ is maximum,

$$|\bar{r}_t^* \bar{e}_i| = \|\bar{r}_t\|_\infty .$$

Now

$$n\|\bar{r}_t\|_\infty^2 = n \left[\max_{1 \leq i \leq n} |r_t^i|^2\right],$$

where r_t^i is the i^{th} component of $\bar{r}_t$. But

$$n \left[\max_{1 \leq i \leq n} |r_t^i|^2\right] \geq \sum_{i=1}^{n} |r_t^i|^2 = \|\bar{r}_t\|^2 .$$

Therefore

$$n\|\bar{r}_t\|_\infty^2 \geq \|\bar{r}_t\|^2$$

so that

$$|\bar{r}_t^* \bar{e}_i| \geq \frac{\|\bar{r}_t\|}{\sqrt{n}} ,$$

since $|\bar{r}_t^* \bar{e}_i| = \|\bar{r}_t\|_\infty$. Also,

$$\bar{e}_i^* A A^* \bar{e}_i = \|A^* \bar{e}_i\|^2 \leq \|A^*\|^2 \, \|\bar{e}_i\|^2 .$$

But $\|A^*\| = \|A\|$, and $\|\bar{e}_i\| = 1$. Thus,

$$\bar{e}_i^* A A^* \bar{e}_i \leq \|A\|^2 .$$

But $\bar{s}_t = A^{-1}\bar{r}_t$ so that

$$\|\bar{s}_t\| = \|A^{-1}\bar{r}_t\| \le \|A^{-1}\| \; \|\bar{r}_t\| ,$$

or

$$\|\bar{r}_t\| \ge \frac{\|\bar{s}_t\|}{\|A^{-1}\|} .$$

From Corollary 3.1.5 one has

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} = 1 - \frac{|\bar{r}_t^*\bar{e}_i|^2}{\|\bar{s}_t\|^2 \; \bar{e}_i^*AA^*\bar{e}_i} .$$

Thus, using the above results,

$$\frac{\|\bar{s}_{t+1}\|^2}{\|\bar{s}_t\|^2} \le 1 - \frac{\dfrac{\|\bar{r}_t\|^2}{n}}{\|\bar{s}_t\|^2 \; \|A\|^2} \le 1 - \frac{\|\bar{s}_t\|^2}{n\|A^{-1}\|^2 \; \|\bar{s}_t\|^2 \; \|A\|^2} .$$

Therefore, since $k(A) = \|A^{-1}\| \; \|A\|$,

$$\|\bar{s}_{t+1}\|^2 \le \|\bar{s}_t\|^2 \left[1 - \frac{1}{n(k(A))^2}\right].$$

<u>Corollary 3.5.2</u>: Consider the scheme (3.7) associated with the system $A\bar{x} = \bar{b}$, where $\bar{v}_t$ is chosen so that $|\bar{v}_t| = \bar{e}$, where $\bar{e}$ is a vector with all its elements one and

$$\bar{v}_t^*\bar{r}_t = \|\bar{r}_t\|_1 = \sum_{i=1}^{n} |r_t^i| ,$$

where r_t^i is the i^{th} component of $\bar{r}_t$. Then

$$\|\bar{s}_{t+1}\|^2 \leq \|\bar{s}_t\|^2 \left[1 - \frac{1}{n(k(A))^2}\right].$$

(Gastinel [2]).

Proof: Using an elementary inequality along with the hypothesis, one has

$$(\bar{v}_t^*\bar{r}_t)^2 = \|\bar{r}_t\|_1^2 = \left[\sum_{i=1}^{n} |r_t^i|\right]\left[\sum_{i=1}^{n} |r_t^i|\right] \geq \sum_{i=1}^{n} |r_t^i|^2$$

or, recalling the definition of $\|\bar{r}_t\|$,

$$(\bar{v}_t^*\bar{r}_t)^2 \geq \sum_{i=1}^{n} |r_t^i|^2 = \|\bar{r}_t\|^2 .$$

Now $\bar{s}_t = A^{-1}\bar{r}_t$ so that

$$\|\bar{s}_t\| \leq \|A^{-1}\| \; \|\bar{r}_t\| .$$

Thus,

$$\bar{v}_t^*\bar{r}_t \geq \|\bar{r}_t\| \geq \frac{\|\bar{s}_t\|}{\|A^{-1}\|} .$$

Also since $\|A^*\| = \|A\|$,

$$\bar{v}_t^*AA^*\bar{v}_t = \|A^*\bar{v}_t\|^2 \leq \|A^*\|^2 \; \|\bar{v}_t\|^2 = \|A\|^2 \; \|\bar{v}_t\|^2 .$$

But

$$\|\bar{v}_t\|^2 = \sum_{i=1}^{n} |v_t^i|^2 = \sum_{i=1}^{n} 1^2 = n .$$

$$v_t^i r_t^i = \frac{|r_t^i|}{r_t^i} r_t^i = |r_t^i|$$

if $r_t^i \neq 0$, and

$$v_t^i r_t^i = 1 r_t^i = |r_t^i|$$

if $r_t^i = 0$, for each i, $1 \leq i \leq n$. Thus,

$$\bar{v}_t^* \bar{r}_t = \sum_{i=1}^{n} |r_t^i| = \|\bar{r}_t\|_1 .$$

<u>Corollary 3.5.3</u>: Consider the scheme (3.7) associated with the system $A\bar{x} = \bar{b}$, where $\bar{v}_t$ is chosen equal to $\bar{r}_t$, the residual vector. Then at each stage

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \cos \theta ,$$

where

$$\cos \theta = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} ,$$

where $k(A)$ is the condition number of the matrix A.

Proof: Note that Theorem 3.5 applies with $V_t = \bar{r}_t$. Thus $\bar{r}_t$ is contained in the space of V_t and from Theorem 3.5,

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Let θ be such that $0 \leq \theta \leq \frac{\pi}{2}$ and

$$\cos\theta = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Then

$$\|\bar{s}_{t+1}\| \leq \|\bar{s}_t\| \cos\theta .$$

Corollary 3.5.4: Consider the scheme (3.7) associated with the system $A\bar{x} = \bar{b}$, where $\bar{v}_t$ is chosen equal to $\bar{r}_t$, the residual vector. Then

(i) $\|\bar{s}_t\| \leq \|\bar{s}_0\| \cos^t\theta$,

(ii) $\|\bar{r}_t\| \leq k(A)\|\bar{r}_t\| \cos^t\theta = \|\bar{r}_0\| \cot\frac{\theta}{2} \cos^t\theta$,

and (iii) $\|\bar{r}_{t+1}\| \leq \|\bar{r}_t\| \cot\theta$,

where θ is such that $0 \leq \theta \leq \frac{\pi}{2}$ and

$$\cos\theta = \frac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}} .$$

Proof: (i) This result will be proved by induction. For $t = 1$, by Corollary 3.5.3,

$$\|\bar{s}_1\| \leq \|\bar{s}_0\| \cos\theta .$$

Now the result is assumed for $t = k$. For $t = k+1$, using Corollary 3.5.3,

$$\|\bar{s}_{k+1}\| \leq \|\bar{s}_k\| \cos\theta .$$

Applying the induction hypothesis yields

$$\|\bar{s}_{k+1}\| \leq \|\bar{s}_0\| \cos^k\theta \cos\theta = \|\bar{s}_0\| \cos^{k+1}\theta \ .$$

Hence the induction is complete and conclusion (i) holds.

(ii) Note that $\bar{r}_t = \bar{b} - A\bar{x}_t = A\bar{s}_t$. Thus

$$\|\bar{r}_t\| = \|A\bar{s}_t\| \leq \|A\| \, \|\bar{s}_t\| \ .$$

Also from the above with $t = 0$, $\bar{s}_0 = A^{-1}\bar{r}_0$. Thus

$$\|\bar{s}_0\| = \|A^{-1}\bar{r}_0\| \leq \|A^{-1}\| \, \|\bar{r}_0\| \ .$$

Therefore, using the result of (i),

$$\|\bar{r}_t\| \leq \|A\| \, \|\bar{s}_t\| \leq \|A\| \, \|\bar{s}_0\| \cos^t\theta \leq \|A\| \, \|A^{-1}\| \, \|\bar{r}_0\| \cos^t\theta \ .$$

But $k(A) = \|A\| \, \|A^{-1}\|$, so that

$$\|\bar{r}_t\| \leq k(A)\|\bar{r}_0\| \cos^t\theta \ .$$

Now

$$\cot^2\frac{\theta}{2} = \frac{\cos^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}} = \frac{1+\cos\theta}{1-\cos\theta} = \frac{1 + \dfrac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}}}{1 - \dfrac{k(A) - \frac{1}{k(A)}}{k(A) + \frac{1}{k(A)}}} \ .$$

Simplifying gives

$$\cot^2\frac{\theta}{2} = \frac{\left[k(A) + \frac{1}{k(A)} + k(A) - \frac{1}{k(A)}\right]\left[k(A) + \frac{1}{k(A)}\right]}{\left[k(A) + \frac{1}{k(A)}\right]\left[k(A) + \frac{1}{k(A)} - k(A) + \frac{1}{k(A)}\right]} \ ,$$

or

$$\cot^2 \frac{\theta}{2} = \frac{2\ k(A)}{\frac{2}{k(A)}} = (k(A))^2 \ .$$

Thus

$$\cot \frac{\theta}{2} = k(A) \ .$$

Therefore,

$$\|\bar{r}_t\| \leq \|\bar{r}_0\| \cot \frac{\theta}{2} \cos^t \theta \ .$$

(iii) From (3.7) the scheme under consideration here is

$$\bar{x}_{t+1} = \bar{x}_t + \frac{\bar{r}_t^* \bar{r}_t A^* \bar{r}_t}{\bar{r}_t^* A A^* \bar{r}_t} \ .$$

$$A\bar{x}_{t+1} = A\bar{x}_t + \frac{\bar{r}_t^* \bar{r}_t A A^* \bar{r}_t}{\bar{r}_t^* A A^* \bar{r}_t}$$

$$\bar{r}_{t+1} = \bar{b} - A\bar{x}_{t+1} = (\bar{b} - A\bar{x}_t) - \frac{\bar{r}_t^* \bar{r}_t A A^* \bar{r}_t}{\bar{r}_t^* A A^* \bar{r}_t}$$

$$\bar{r}_{t+1} = \bar{r}_t - \frac{\|\bar{r}_t\|^2 \ A A^* \bar{r}_t}{\bar{r}_t^* A A^* \bar{r}_t}$$

$$\|\bar{r}_{t+1}\|^2 = \bar{r}_{t+1}^* \bar{r}_{t+1} = \left[\bar{r}_t^* - \frac{\|\bar{r}_t\|^2 (A A^* \bar{r}_t)^*}{\bar{r}_t^* A A^* \bar{r}_t} \right] \left[\bar{r}_t - \frac{\|\bar{r}_t\|^2 A A^* \bar{r}_t}{\bar{r}_t^* A A^* \bar{r}_t} \right]$$

$$\|\bar{r}_{t+1}\|^2 = \bar{r}_t^*\bar{r}_t - \frac{\|\bar{r}_t\|^2\bar{r}_t^*AA^*\bar{r}_t}{\bar{r}_t^*AA^*\bar{r}_t} - \frac{\|\bar{r}_t\|^2\bar{r}_t^*AA^*\bar{r}_t}{\bar{r}_t^*AA^*\bar{r}_t}$$

$$+ \frac{\|\bar{r}_t\|^4\bar{r}_t^*AA^*AA^*\bar{r}_t}{(\bar{r}_t^*AA^*\bar{r}_t)^2}$$

$$\|\bar{r}_{t+1}\|^2 = \|\bar{r}_t\|^2 - \|\bar{r}_t\|^2 - \|\bar{r}_t\|^2 + \frac{\|\bar{r}_t\|^4\bar{r}_t^*(AA^*)^2\bar{r}_t}{(\bar{r}_t^*AA^*\bar{r}_t)^2}$$

Therefore,

$$\frac{\|\bar{r}_{t+1}\|^2}{\|\bar{r}_t\|^2} = \frac{\bar{r}_t^*\bar{r}_t\ \bar{r}_t^*(AA^*)^2\bar{r}_t}{(\bar{r}_t^*AA^*\bar{r}_t)^2} - 1 .$$

Applying the Kantorovich inequality, Theorem 3.4, one has

$$\frac{(\bar{r}_t^*AA^*\bar{r}_t)^2}{\bar{r}_t^*\bar{r}_t\ \bar{r}_t^*(AA^*)^2\bar{r}_t} \geq \sin^2\theta .$$

Therefore,

$$\frac{\|\bar{r}_{t+1}\|^2}{\|\bar{r}_t\|^2} \leq \frac{1}{\sin^2\theta} - 1 = \csc^2\theta - 1 = \cot^2\theta .$$

Thus,

$$\|\bar{r}_{t+1}\|^2 \leq \|\bar{r}_t\|^2 \cot^2\theta ,$$

or

$$\|\bar{r}_{t+1}\| \leq \|\bar{r}_t\| \cot\theta .$$

This completes the proof.

The Case, "A Is Positive Definite."

Now it is assumed that the matrix A, in the equation $A\bar{x} = \bar{b}$, is a positive definite Hermitian matrix. Then the natural choice for G is $G = A$. With this choice the norm used is

$$\|\bar{x}\|_A = (\bar{x}^* A \bar{x})^{1/2} .$$

The remaining theorems deal with the iterative schemes resulting from this choice of G.

Theorem 3.6: Consider the system $A\bar{x} = \bar{b}$, where the matrix A is positive definite and Hermitian. Then the scheme

$$\bar{x}_{t+1} = \bar{x}_t + Y_t(Y_t^* A Y_t)^{-1}\, Y_t^*(\bar{b} - A\bar{x}_t)$$

is a norm reducing projective iterative method. (Recall Y_t is an $n x k_t$, $0 \leq k_t \leq n$, matrix whose columns are linearly independent. Also, the system is norm reducing in the sense of $\|\cdot\|_A$.) Further

$$\|\bar{s}_t\|_A^2 - \|\bar{s}_{t+1}\|_A^2 = \bar{r}_t^* Y_t (Y_t^* A Y_t)^{-1} Y_t^* \bar{r}_t ,$$

where $\bar{r}_t$ is the residual vector.

Proof: Note that $A = A^*$ and $\bar{r}_t = A\bar{s}_t$. Thus,

$$\bar{s}_t^* A Y_t (Y_t^* A Y_t)^{-1}\, Y_t^* A \bar{s}_t = \bar{r}_t^* Y_t (Y_t^* A Y_t)^{-1} Y_t^* \bar{r}_t .$$

Using this fact and applying Theorem 3.1 with $G = A$, one obtains the conclusion of the theorem.

Corollary 3.6.1: Consider the iterative scheme

$$\bar{x}_{t+1} = \bar{x}_t + \frac{\bar{y}_t \bar{y}_t^* (\bar{b} - A\bar{x}_t)}{\bar{y}_t^* A \bar{y}_t} \tag{3.8}$$

associated with the system $A\bar{x} = \bar{b}$, where A is a positive definite Hermitian matrix and $\bar{y}_t$ is chosen to be a natural unit vector $\bar{e}_i$ which maximizes $|\bar{e}_i^* \bar{r}_t|$ over the possible choices of i. Then

$$\|\bar{s}_{t+1}\|_A^2 \leq \|\bar{s}_t\|_A^2 \left[1 - \frac{1}{nk(A)}\right],$$

where $k(A)$ is the condition number of A.

Proof: Let r_t^i denote the i^{th} component of $\bar{r}_t$.

$$\|\bar{r}_t\|^2 = \sum_{j=1}^{n} |r_t^j|^2 \leq \sum_{j=1}^{n} \max_{1 \leq j \leq n} |r_t^j|^2 = n|\bar{e}_i^* \bar{r}_t|^2$$

Thus,

$$|\bar{e}_i^* \bar{r}_t| \geq \frac{\|\bar{r}_t\|}{\sqrt{n}} .$$

Consider the functional $f(\bar{x})$ defined by

$$f(\bar{x}) = \bar{e}_i^* \bar{x} .$$

From properties of linear functionals defined in this way,

$$\|f\| = \|\bar{e}_i\| .$$

But $\|\bar{e}_i\| = 1$, so that $\|f\| = 1$. Now

$$\bar{e}_i^* A\bar{e}_i = |f(A\bar{e}_i)| \le \|f\| \, \|A\bar{e}i\| \le \|A\| \, \|\bar{e}_i\| = \|A\| .$$

Let P be a matrix such that $A = P^*P$. (P exists since A is a positive definite Hermitian matrix.) Then

$$\|\bar{s}_t\|_A = (\bar{s}_t^* A\bar{s}_t)^{1/2} = (\bar{s}_t^* P^* P\bar{s}_t)^{1/2} = \|P\bar{s}_t\|.$$

Since $\bar{s}_t = A^{-1}\bar{r}_t$,

$$\|\bar{s}_t\|_A = \|P\bar{s}_t\| = \|PA^{-1}\bar{r}_t\| = \|P(P^*P)^{-1}\bar{r}_t\| = \|PP^{-1}(P^*)^{-1}\bar{r}_t\| ,$$

or

$$\|\bar{s}_t\|_A = \|(P^*)^{-1}\bar{r}_t\| \le \|(P^*)^{-1}\| \, \|\bar{r}_t\| .$$

But $\|(P^*)^{-1}\| = \|P^{-1}\|$, so that

$$\|\bar{r}_t\| \ge \frac{\|\bar{s}_t\|_A}{\|P^{-1}\|} .$$

Now using Theorem 3.6, where in this case $Y_t = \bar{e}_i$, one has

$$\|\bar{s}_t\|_A^2 - \|\bar{s}_{t+1}\|_A^2 = \frac{\bar{r}_t^* \bar{e}_i \bar{e}_i^* \bar{r}_t}{\bar{e}_i^* A\bar{e}_i} = \frac{|\bar{e}_i^* \bar{r}_t|^2}{\bar{e}_i^* A\bar{e}_i} .$$

Therefore,

$$\|\bar{s}_t\|_A^2 - \|\bar{s}_{t+1}\|_A^2 \ge \frac{\|\bar{r}_t\|^2}{n\|A\|} \ge \frac{\|\bar{s}_t\|_A^2}{n\|P^{-1}\|^2 \|P^*P\|} .$$

But $\|P^*P\| \le \|P^*\| \, \|P\| = \|P\|^2$. Thus

$$\|\bar{s}_t\|_A^2 - \|\bar{s}_{t+1}\|_A^2 \geq \frac{\|\bar{s}_t\|_A^2}{n\|P^{-1}\|^2\|P\|^2} = \frac{\|\bar{s}_t\|_A^2}{n(k(P))^2} ,$$

where $k(P)$ is the condition number of the matrix P. It will now be shown that $k(A) = (k(P))^2$. Let $r_\sigma(A)$ denote the spectral radius of A and $r_\sigma(A^{-1})$ denote the spectral radius of A^{-1}. Then, using properties of $\|\cdot\|$,

$$k(A) = \|A\| \; \|A^{-1}\| = r_\sigma(A) \; r_\sigma(A^{-1}) = r_\sigma(P^*P) r_\sigma((P^*P)^{-1}) ,$$

or

$$k(A) = r_\sigma(P^*P) \; r_\sigma(P^{-1}(P^*)^{-1}) = r_\sigma(P^*P) \; r_\sigma(P^{-1}(P^{-1})^*) .$$

Thus,

$$k(A) = \|P\|^2 \; \|P^{-1}\|^2 = (k(P))^2 .$$

Therefore,

$$\|\bar{s}_{t+1}\|_A^2 \leq \|\bar{s}_t\|_A^2 \left[1 - \frac{1}{n(k(A))}\right] .$$

This completes the proof.

Corollary 3.6.2: Consider the iterative scheme (3.8) associated with the system $A\bar{x} = \bar{b}$, where A is positive definite and Hermitian. The $\bar{y}_t$ are chosen so that $|\bar{y}_t| = \bar{e}$, where $\bar{e}$ is a vector with all its elements equal to one, and

$$\bar{y}_t^* \bar{r}_t = \|\bar{r}_t\|_1 .$$

Then

$$\|\bar{s}_{t+1}\|_A^2 \le \|\bar{s}_t\|_A^2 \left[1 - \frac{1}{n(k(A))}\right],$$

where $k(A)$ is the condition number of the matrix A.

Proof: Let P be a matrix such that $A = P^*P$. (P exists since A is positive definite and Hermitian.) From the proof of Corollary 3.6.1, one has

$$\|\bar{s}_t\|_A \le \|P^{-1}\| \, \|\bar{r}_t\| .$$

Now

$$\|\bar{r}_t\| = \left[\sum_{i=1}^{n} |r_t^i|^2\right]^{1/2} \le \sum_{i=1}^{n} |r_t^i| = \|\bar{r}_t\|_1 ,$$

where r_t^i is the i^{th} component of $\bar{r}_t$. Therefore

$$\|\bar{r}_t\|_1 \ge \frac{\|\bar{s}_t\|_A}{\|P^{-1}\|} .$$

Consider the bounded linear functional

$$f(\bar{x}) = \bar{y}_t^*\bar{x} .$$

From properties of linear functionals defined in this way,

$$\|f\| = \|\bar{y}_t\| .$$

But

$$\|\bar{y}_t\| = \left[\sum_{i=1}^{n} |y_t^i|^2\right]^{1/2} = \left[\sum_{i=1}^{n} 1^2\right]^{1/2} = \sqrt{n} .$$

Thus $\|f\| = \sqrt{n}$. Now

$$\overline{y}_t^* A\overline{y}_t = |f(A\overline{y}_t)| \leq \|f\| \, \|A\overline{y}_t\| \leq \sqrt{n} \, \|A\| \, \|\overline{y}_t\| = \sqrt{n} \, \|A\| \sqrt{n} .$$

Therefore,

$$\overline{y}_t^* A\overline{y}_t \leq n\|A\| .$$

Next, applying Theorem 3.6 with $Y_t = \overline{y}_t$ yields

$$\|\overline{s}_t\|_A^2 - \|\overline{s}_{t+1}\|_A^2 = \frac{\overline{r}_t^* \overline{y}_t \overline{y}_t^* \overline{r}_t}{\overline{y}_t^* A\overline{y}_t} = \frac{\|\overline{r}_t^* \overline{y}_t\|^2}{\overline{y}_t^* A\overline{y}_t} .$$

Thus

$$\|\overline{s}_t\|_A^2 - \|\overline{s}_{t+1}\|_A^2 \geq \frac{\|\overline{r}_t\|_1^2}{n\|A\|} \geq \frac{\|\overline{s}_t\|_A^2}{n\|P^*P\| \, \|P^{-1}\|^2} \geq \frac{\|\overline{s}_t\|_A^2}{n\|P^*\| \, \|P\| \, \|P^{-1}\|^2} .$$

Therefore,

$$\|\overline{s}_{t+1}\|_A^2 \leq \|\overline{s}_t\|_A^2 \left[1 - \frac{1}{n\|P\|^2\|P^{-1}\|^2}\right] .$$

Now recall from the proof of Corollary 3.6.1 that

$$\|P\|^2 \, \|P^{-1}\|^2 = (k(P))^2 = k(A) .$$

Thus,

$$\|\overline{s}_{t+1}\|_A^2 \leq \|\overline{s}_t\|_A^2 \left[1 - \frac{1}{n(k(A))}\right] .$$

<u>Corollary 3.6.3</u>: Consider the iterative scheme (3.8) associated with

the system $A\bar{x} = \bar{b}$, where A is positive definite and Hermitian. If $\bar{y}_t$ is chosen equal to $\bar{r}_t$ at each stage, then

$$\|\bar{s}_{t+1}\|_A \leq \|\bar{s}_t\|_A \cos\theta ,$$

where θ is such that $0 \leq \theta \leq \frac{\pi}{2}$, and

$$\cos\theta = \frac{q - \frac{1}{q}}{q + \frac{1}{q}} ,$$

where $q = (k(A))^{1/2}$.

Proof: From Theorem 3.6 with $Y_t = \bar{r}_t$ one has

$$\|\bar{s}_{t+1}\|_A^2 = \|\bar{s}_t\|_A^2 \left[1 - \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*\bar{r}_t}{\bar{s}_t^*A\bar{s}_t \;\; \bar{r}_t^*A\bar{r}_t} \right]$$

using the definition of $\|\bar{s}_t\|_A$. Let P be a matrix such that $A = P^*P$. Then since $\bar{s}_t = A^{-1}\bar{r}_t$,

$$\frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*\bar{r}_t}{\bar{s}_t^*A\bar{s}_t \;\; \bar{r}_t^*A\bar{r}_t} = \frac{|\bar{r}_t^*\bar{r}_t|^2}{\bar{r}_t^*(A^{-1})^*AA^{-1}\bar{r}_t\bar{r}_t^*A\bar{r}_t}$$

$$= \frac{|\bar{r}_t^*\bar{r}_t|^2}{\bar{r}_t^*(P^*P)^{-1}\bar{r}_t\bar{r}_t^*(P^*P)\bar{r}_t} .$$

Now from the proof of Corollary 3.6.1, $k(P) = (k(A))^{1/2} = q$. Thus applying the inequality of Kantorovich,

$$\frac{|\bar{r}_t^*\bar{r}_t|^2}{\bar{r}_t^*(P^*P)^{-1}\bar{r}_t\bar{r}_t^*(P^*P)\bar{r}_t} \geq \sin^2\theta' ,$$

where $0 \le \theta' \le \frac{\pi}{2}$, and

$$\cos \theta' = \frac{k(P) - \frac{1}{k(P)}}{k(P) + \frac{1}{k(P)}} = \frac{q - \frac{1}{q}}{q + \frac{1}{q}} .$$

Note that $\cos \theta = \cos \theta'$ and both θ and θ' lie in the first quadrant. Thus $\theta = \theta'$. Therefore,

$$\|\bar{s}_{t+1}\|_A^2 \le \|\bar{s}_t\|_A^2 (1 - \sin^2 \theta) = \|\bar{s}_t\|_A^2 \cos^2 \theta ,$$

or

$$\|\bar{s}_{t+1}\|_A \le \|\bar{s}_t\|_A \cos \theta .$$

<u>Corollary 3.6.4</u>: Consider the iterative scheme (3.8) associated with the system $A\bar{x} = \bar{b}$, where A is positive definite and Hermitian. Choose $\bar{y}_t = \bar{r}_t$ and let θ be defined as in Corollary 3.6.3. Then

(i) $\|\bar{s}_t\| \le \|\bar{s}_0\| \cot \frac{\theta}{2} \cos^t \theta$;

(ii) $\|\bar{r}_t\| \le \|\bar{r}_0\| \cot \frac{\theta}{2} \cos^t \theta$;

and (iii) $\|\bar{r}_{t+1}\| \le \|\bar{r}_t\| \cos \theta$.

Proof: (i) First it is necessary to establish that

$$\|\bar{s}_t\|_A \le \|\bar{s}_0\|_A \cos^t \theta .$$

This will be done by induction. For $t = 1$, applying Corollary 3.6.3 yields

$$\|\bar{s}_1\|_A \le \|\bar{s}_0\|_A \cos\theta .$$

Now the result is assumed for $t = k$. Then for $t = k+1$ one has, applying Corollary 3.6.3,

$$\|\bar{s}_{k+1}\|_A \le \|\bar{s}_k\|_A \cos\theta \le \|\bar{s}_0\|_A \cos^k\theta \cos\theta = \|\bar{s}_0\|_A \cos^{k+1}\theta .$$

This completes the induction and the result is established. Let P be a matrix such that $A = P^*P$. Then

$$\|\bar{s}_0\|_A = (\bar{s}_0^* A\bar{s}_0)^{1/2} = (\bar{s}_0^* P^* P\bar{s}_0)^{1/2} = \|P\bar{s}_0\| \le \|P\|\|\bar{s}_0\|$$

and

$$\|\bar{s}_t\| = \|P^{-1}P\bar{s}_t\| \le \|P^{-1}\| \|P\bar{s}_t\| = \|P^{-1}\|(\bar{s}_t^* P^* P\bar{s}_t)^{1/2}$$

$$\le \|P^{-1}\|(\bar{s}_t^* A\bar{s}_t)^{1/2} = \|P^{-1}\| \|\bar{s}_t\|_A .$$

Therefore,

$$\frac{\|\bar{s}_t\|}{\|P^{-1}\|} \le \|\bar{s}_t\|_A \le \|\bar{s}_0\|_A \cos^t\theta \le \|P\| \|\bar{s}_0\| \cos^t\theta ,$$

or

$$\|\bar{s}_t\| \le \|P\| \|P^{-1}\| \|\bar{s}_0\| \cos^t\theta .$$

Now from the proof of Corollary 3.6.1,

$$k(P) = \|P\| \|P^{-1}\| = (k(A))^{1/2} = q .$$

$$\cot^2 \frac{\theta}{2} = \frac{1 + \cos\theta}{1 - \cos\theta} = \frac{\left[q + \frac{1}{q} + q - \frac{1}{q}\right]\left[q + \frac{1}{q}\right]}{\left[q + \frac{1}{q}\right]\left[q + \frac{1}{q} - q + \frac{1}{q}\right]} = \frac{2q}{\frac{2}{q}} = q^2$$

$$\|P\| \, \|P^{-1}\| = q = \cot \frac{\theta}{2}$$

Therefore,

$$\|\bar{s}_t\| \le \|\bar{s}_0\| \cot \frac{\theta}{2} \cos^t \theta .$$

This proves (i).

(ii) Using the same P as in the proof of (i),

$$\|\bar{s}_0\|_A = \|P\bar{s}_0\| = \|PA^{-1}\bar{r}_0\| = \|PP^{-1}(P^*)^{-1}\bar{r}_0\| = \|(P^{-1})^*\bar{r}_0\|$$

$$\le \|(P^{-1})^*\| \, \|\bar{r}_0\| = \|P^{-1}\| \, \|\bar{r}_0\|$$

and

$$\|\bar{r}_t\| = \|A\bar{s}_t\| = \|P^*P\bar{s}_t\| \le \|P^*\| \, \|P\bar{s}_t\| = \|P\| \, \|\bar{s}_t\|_A .$$

Thus,

$$\frac{\|\bar{r}_t\|}{\|P\|} \le \|\bar{s}_t\|_A \le \|\bar{s}_0\|_A \cos^t\theta \le \|P^{-1}\| \, \|\bar{r}_0\| \cos^t \theta ,$$

or

$$\|\bar{r}_t\| \le \|P\| \, \|P^{-1}\| \, \|\bar{r}_0\| \cos^t\theta = \|\bar{r}_0\| \cot \frac{\theta}{2} \cos^t \theta .$$

This proves (ii).

(iii) From the definition of the iteration (3.8), one has for $t \ge 0$,

$$\bar{x}_{t+1} = \bar{x}_t + \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} .$$

Thus,

$$\bar{r}_{t+1} = \bar{b} - A\bar{x}_{t+1} = \bar{b} - A\bar{x}_t - \frac{\bar{r}_t^*\bar{r}_tA\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} = \bar{r}_t - \frac{\bar{r}_t^*\bar{r}_tA\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} .$$

Now, since A is Hermitian,

$$\bar{r}_{t+1}^*\bar{r}_{t+1} = \left[\bar{r}_t^* - \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A^*}{\bar{r}_t^*A\bar{r}_t}\right]\left[\bar{r}_t - \frac{\bar{r}_t^*\bar{r}_tA\bar{r}_t}{\bar{r}_t^*A\bar{r}_t}\right]$$

$$= \bar{r}_t^*\bar{r}_t - \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} - \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A^*\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} + \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*\bar{r}_t\bar{r}_t^*A^*A\bar{r}_t}{\bar{r}_t^*A\bar{r}_t\ \bar{r}_t^*A\bar{r}_t}$$

$$= \bar{r}_t^*\bar{r}_t - \bar{r}_t^*\bar{r}_t - \frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A\bar{r}_t}{\bar{r}_t^*A\bar{r}_t} + \frac{(\bar{r}_t^*\bar{r}_t)^2\bar{r}_t^*A^2\bar{r}_t}{(\bar{r}_t^*A\bar{r}_t)^2}$$

$$= (\bar{r}_t^*\bar{r}_t)\left[\frac{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A^2\bar{r}_t}{(\bar{r}_t^*A\bar{r}_t)^2} - 1\right] = \|\bar{r}_{t+1}\|^2 .$$

Let P be the same as in the proof of (i) and consider

$$\frac{(\bar{r}_t^*A\bar{r}_t)^2}{\bar{r}_t^*\bar{r}_t\bar{r}_t^*A^2\bar{r}_t} = \frac{(\bar{r}_t^*P^*P\bar{r}_t)^2}{\bar{r}_t^*\bar{r}_t\bar{r}_t^*(P^*P)^2\bar{r}_t} .$$

Applying the Kantorovich inequality, Theorem 3.4, with $\bar{x} = P\bar{r}_t$ yields

$$\frac{(\bar{r}_t^* A \bar{r}_t)^2}{\bar{r}_t^* \bar{r}_t \bar{r}_t^* A^2 \bar{r}_t} \geq \sin^2 \theta \; .$$

Therefore,

$$\|\bar{r}_{t+1}\|^2 \leq \|\bar{r}_t\|^2 \left[\frac{1}{\sin^2 \theta} - 1\right] = \|\bar{r}_t\|^2 \,(\csc^2\theta - 1) = \|\bar{r}_t\|^2 \cot^2 \theta$$

or

$$\|\bar{r}_{t+1}\| \leq \|\bar{r}_t\| \cot \theta \; .$$

This completes the proof.

BIBLIOGRAPHY

[1] Faddeev, D. K., and Faddeeva, V. N., Computational Methods of Linear Algebra, San Francisco and London, W. H. Freeman and Company, 1963, pp. 385-405.

[2] Gastinel, Noël, "Procédé Iteratif pour la Résolution Numérique d'un Système d'Équations Linéaires," Comptes Rendus de l'Académie des Sciences Paris, Vol. 246, (1958), pp. 2571-2574.

[3] Hestenes, M. R., and Stiefel, Eduard, "Methods of Conjugate Gradients for Solving Linear Systems," Journal of Research of the National Bureau of Standards, Vol. 49 (1952), pp. 409-436.

[4] Householder, A. S., "On Certain Iterative Methods for Solving Linear Systems," Numerische Mathematik, Vol. 2 (1960), pp. 55-59.

[5] ________, Principles of Numerical Analysis, New York, McGraw-Hill, 1953, pp. 44-52.

[6] ________, "The Kantorovich and Some Related Inequalities," SIAM Review, Vol. 7 (1965), pp. 463-473.

[7] ________, The Theory of Matrices in Numerical Analysis, New York, Blaisdell Publishing Company, 1964, pp. 81-84, 98-103.