Spectral characterization for von Neumann’s iterative algorithm in $\mathbb{R}^n$

Rudy Joly, Marco López, Douglas Mupasiri and Michael Newsome
Spectral characterization for von Neumann’s iterative algorithm in $\mathbb{R}^n$

Rudy Joly, Marco López, Douglas Mupasiri and Michael Newsome

(Communicated by Jim Haglund)

Our work is motivated by a theorem proved by von Neumann: Let $S_1$ and $S_2$ be subspaces of a closed Hilbert space $X$ and let $x \in X$. Then

$$\lim_{k \to \infty} (P_{S_2}P_{S_1})^k(x) = P_{S_1 \cap S_2}(x),$$

where $P_S$ denotes the orthogonal projection of $x$ onto the subspace $S$. We look at the linear algebra realization of the von Neumann theorem in $\mathbb{R}^n$. The matrix $A$ that represents the composition $P_{S_2}P_{S_1}$ has a form simple enough that the calculation of $\lim_{k \to \infty} A^k x$ becomes easy. However, a more interesting result lies in the analysis of eigenvalues and eigenvectors of $A$ and their geometrical interpretation. A characterization of such eigenvalues and eigenvectors is shown for subspaces with dimension $n - 1$.

1. Introduction

In Euclidean $n$-space, we wish to find the point $x_\infty$ in the intersection of two $(n - 1)$-dimensional subspaces, $S_1$ and $S_2$, that is closest to an initial point $x_0$ in $\mathbb{R}^n$. That is, we want $x_\infty \in S_1 \cap S_2$ to be such that

$$\|x_0 - x_\infty\| \leq \|x_0 - y\| \text{ for all } y \in S_1 \cap S_2.$$

We call $x_\infty$ the orthogonal projection of $x_0$ onto $S_1 \cap S_2$. We start by stating von Neumann’s theorem; see [Deutsch 2001], for example.

**Theorem 1.** Let $S_1$ and $S_2$ be subspaces of a closed Hilbert space $X$ and let $x \in X$. Then

$$\lim_{k \to \infty} (P_{S_2}P_{S_1})^k(x) = P_{S_1 \cap S_2}(x),$$

where $P_S$ denotes the orthogonal projection onto the subspace $S$.

Von Neumann’s theorem provides an iterative procedure (left-hand side of (1-1)) to find the orthogonal projection of $x$ onto $S_1 \cap S_2$ (right-hand side of (1-1)).

**MSC2010:** primary 41A65; secondary 47N10.

**Keywords:** orthogonal projections, von Neumann, best approximations.
2. An example in $\mathbb{R}^2$

To illustrate von Neumann’s theorem we consider the $\mathbb{R}^2$ case. Let $a_1, b_1, a_2, b_2 \in \mathbb{R}$ and let

$$S_1 = \{(x, y) \mid a_1 x + b_1 y = 0\} \quad \text{and} \quad S_2 = \{(x, y) \mid a_2 x + b_2 y = 0\}.$$ 

In order for $S_1$ and $S_2$ to be distinct 1-dimensional subspaces, we require that the $a_i$ and $b_i$ are not both zero\(^1\) and that $a_1/b_1 \neq a_2/b_2$. Since the orthogonal projection onto a subspace is a linear transformation, we can represent such transformations by matrices. In the plane, the matrix that projects any point in $\mathbb{R}^2$ onto $S_i$ is given by

$$A_i = \frac{1}{a_i^2 + b_i^2} \begin{pmatrix} b_i^2 & -a_i b_i \\ -a_i b_i & a_i^2 \end{pmatrix},$$

where $i = 1, 2$. Therefore, the matrix $A = A_2 A_1$ gives us the composition of the two projections.

$$A = \frac{a_1 a_2 + b_1 b_2}{(a_1^2 + b_1^2)(a_2^2 + b_2^2)} \begin{pmatrix} b_1 b_2 & -a_1 b_2 \\ -a_2 b_1 & a_1 a_2 \end{pmatrix}$$

To compute iterations of the matrix $A$, we wish to express $A$ in terms of a diagonal matrix $D$ similar to $A$. This is possible, of course, if $A$ is nondefective; that is, if the dimension of each of the eigenspaces of $A$ is equal to the multiplicity of the corresponding eigenvalue. It is easily shown that $A$ is nondefective in the $\mathbb{R}^2$ case. The matrix $S$ of eigenvectors of $A$ is then

$$S = \begin{pmatrix} a_1 & b_1 \\ b_1 & -a_2 \end{pmatrix},$$

with $D$ being

$$D = S^{-1} A S.$$

Computing powers of the matrix $A$ is then a matter of raising the eigenvalues of $A$ to that power:

$$A^k = S D^k S^{-1}.$$  

Applying von Neumann’s theorem to this equation, we obtain

$$\lim_{k \to \infty} (A_2 A_1)^k = \lim_{k \to \infty} A^k = S \left( \lim_{k \to \infty} D^k \right) S^{-1} = A_\infty,$$

where $A_\infty$ is the matrix representation of $P_{S_1 \cap S_2}$. Note that the limit exists if the eigenvalues of $A$ have absolute value less than or equal to unity.

\(^1\)If, say, $a_1 = b_1 = 0$ then $S_1 = \mathbb{R}^2$. 

3. Solution algorithm

It is possible to extend the solution method in the previous section to \( \mathbb{R}^n \). Here we present a brief outline of the solution algorithm, as explained in [Hoffman and Kunze 1971].

(1) Choose bases for \( S_1 \) and \( S_2 \).

(2) Use the Gram–Schmidt procedure to produce orthonormal bases \( \beta^{(1)} \) and \( \beta^{(2)} \) for \( S_1 \) and \( S_2 \) respectively:

\[
\beta^{(1)} = \{ u^{(1)}_1, \ldots, u^{(1)}_{n-1} \}, \quad \beta^{(2)} = \{ u^{(2)}_1, \ldots, u^{(2)}_{n-1} \}.
\]

(3) Use the standard basis \( \beta = \{ e_1, \ldots, e_n \} \) for the parent vector space \( \mathbb{R}^n \).

(4) Use the following general formula to obtain the matrix representations \( A_i \), with \( i = 1, 2 \), of the orthogonal projections \( P_i : \mathbb{R}^n \to S_i \):

\[
A_i = \left[ \left( \sum_{j=1}^{n-1} \langle e_1, u^{(i)}_j \rangle u^{(i)}_j \right), \ldots, \left( \sum_{j=1}^{n-1} \langle e_n, u^{(i)}_j \rangle u^{(i)}_j \right) \right].
\]

(5) Compute \( A = A_2 A_1 \). Find the eigenvalues \( \lambda_1, \ldots, \lambda_n \) and corresponding independent eigenvectors \( E_1, \ldots, E_n \) of \( A \). These give us the \( n \times n \) matrices

\[
D = \begin{pmatrix} \lambda_1 & 0 \\ \vdots & \ddots \\ 0 & \cdots & \lambda_n \end{pmatrix}, \quad S = (E_1, \ldots, E_n).
\]

(6) Compute \( S^{-1} \).

(7) Iteration now proceeds as follows:

\[
v_k = A v_{k-1} = (SDS^{-1}) v_{k-1} = (SDS^{-1})(SDS^{-1}) v_{k-2} = \cdots = (SD^k S^{-1}) v_0 = A^k v_0
\]

for \( k = 1, 2, 3, \ldots \).

(8) Finally, we obtain \( v_\infty = [S(\lim_{k \to \infty} D^k) S^{-1}] v_0 \).

In step (5), we rely on the assumption that the matrix \( A \) is nondefective in order to find a similar diagonal matrix. We address this question in Section 5.

4. Eigenvalues in \( \mathbb{R}^3 \): geometric argument

If we consider two 2-dimensional subspaces in 3-space, \( S_1 \) and \( S_2 \), it is easy to illustrate geometrically the eigenvectors of the alternating projections. By examining a picture of two planes containing the origin in \( \mathbb{R}^3 \), we see three different types of eigenvectors; the first two are trivial, but the third is less so (refer to Figure 1).
Figure 1. Top left: a vector orthogonal to $S_1$ gets projected to the origin (eigenvalue 0). Top right: a vector in $S_1 \cap S_2$ remains fixed (eigenvalue 1). Bottom: a vector in $(S_1 \cap S_2)^\perp$ gets projected to a collinear vector (eigenvalue in $[0, 1]$).

(1) A vector orthogonal to $S_1$ is in the kernel of $P_{S_1}$; therefore, it is an eigenvector of $P_{S_1}$ with eigenvalue 0.

(2) A vector in $S_1 \cap S_2$ is an eigenvector of both $P_{S_2}$ and $P_{S_1}$ with eigenvalue 1.

(3) A vector in the orthogonal complement $(S_1 \cap S_2)^\perp$ will stay in $(S_1 \cap S_2)^\perp$ as it is projected orthogonally onto $S_1$ and $S_2$; i.e., $(S_1 \cap S_2)^\perp$ is invariant under both $P_{S_1}$ and $P_{S_2}$. Therefore, a vector in $S_2 \cap (S_1 \cap S_2)^\perp$ is an eigenvector of $P_{S_2}P_{S_1}$. We claim that this eigenvector corresponds to an eigenvalue in the interval $[0, 1]$.

It is easy to see from this geometric argument the characterization of eigenvalues in the case of $\mathbb{R}^3$. Next we address the question of whether this geometric intuition somehow generalizes to $\mathbb{R}^n$.

5. Characterization of eigenvalues in $\mathbb{R}^n$.

When we consider $(n - 1)$-dimensional subspaces in $\mathbb{R}^n$, it is easy to see that the first two eigenvectors described in Section 4 generalize to higher dimensions. It
is less trivial to show that the third type of eigenvector also generalizes to higher dimensions, and that these three types of vectors fully characterize the spectrum of $P_{S_2} P_{S_1}$.

Let $S_1$ and $S_2$ be $(n - 1)$-dimensional subspaces of $\mathbb{R}^n$ with $S_1 \neq S_2$.

**Lemma 2.** $S_1 \cap S_2$ is a proper subspace of $\mathbb{R}^n$ with $\dim(S_1 \cap S_2) = n - 2$.

**Proof.** The intersection of two subspaces is always a subspace. Note that for two distinct subspaces, we have

$$n = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2).$$

Therefore,

$$\dim(S_1 \cap S_2) = \dim(S_1) + \dim(S_2) - n = n - 1 + n - 1 - n = n - 2. \quad \square$$

Now, let $S_3 = (S_1 \cap S_2) \perp$. Note that $n = \dim(S_1 \cap S_2) + \dim(S_3)$, which implies that $\dim(S_3) = 2$.

**Lemma 3.** $\dim(S_3 \cap S_1) = \dim(S_3 \cap S_2) = 1$.

**Proof.** We write $\dim(S_3 \cap S_1) = \dim(S_3) + \dim(S_1) - n = 2 + n - 1 - n = 1$. Similarly, $\dim(S_3 \cap S_2) = 1$. \quad \square

**Lemma 4.** Let $T_1 : \mathbb{R}^n \to S_1$ and $T_2 : \mathbb{R}^n \to S_2$ be the orthogonal projections onto $S_1$ and $S_2$, respectively. Then $S_3$ is invariant under $T_1$ and $T_2$.

**Proof.** Let $\{w, w^\perp\}$ be a basis for $S_3$ such that $w \in S_1$ and $w^\perp \in S_1^\perp$. If $v_0 \in S_3$, then $v_0 = c_1 w + c_2 w^\perp$ for some scalars $c_1, c_2$; therefore,

$$T_1(v_0) = c_1 T_1(w) + c_2 T_1(w^\perp) = c_1 w \in S_3.$$  

Similarly, we can construct a basis $\{u, u^\perp\}$ for $S_3$ such that $u \in S_2$ and $u^\perp \in S_2^\perp$ to conclude that $T_2(v_0) \in S_3$. \quad \square

Now we are ready to prove the following theorem. Let $\theta$ be the angle between two hyperplanes defined as the angle between two vectors $n_1$ and $n_2$ normal to $S_1$ and $S_2$, respectively. Note that $n_1, n_2 \in S_3$.

**Theorem 5.** Let $S_1$ and $S_2$ be distinct $(n - 1)$-dimensional subspaces of $\mathbb{R}^n$, and let $T_1 : \mathbb{R}^n \to S_1$ and $T_2 : \mathbb{R}^n \to S_2$ be the orthogonal projections onto $S_1$ and $S_2$, respectively. Also, let $0 < \theta < \frac{\pi}{2}$ be the angle between the two hyperplanes. The spectrum of $T := T_2 T_1$ is characterized by the following eigenvalues and multiplicities:

$$\lambda_1 = 0, \ m_1 = 1, \ \lambda_2 = 1, \ m_2 = n - 2, \ \lambda_3 = \cos^2 \theta, \ m_3 = 1.$$
Proof. First, consider \( u_0 \) to be a vector orthogonal to \( S_1 \). Then \( T(u_0) = 0 \), and so \( m_1 \geq 1 \). Now let \( \{w_1, \ldots, w_{n-2}\} \) be a basis for \( S_1 \cap S_2 \). Then \( T(w_i) = w_i \) for all \( 1 \leq i \leq n-2 \). Therefore, \( \lambda_2 = 1 \) is an eigenvalue. Since the basis vectors for \( S_1 \cap S_2 \) are linearly independent eigenvectors corresponding to \( \lambda_2 \), we have \( m_2 \geq n - 2 \). Furthermore, consider \( v_0 \in S_3 \cap S_2 \). Then \( T(v_0) \in S_3 \) by Lemma 4, and \( T(v_0) \in S_2 \) since the range of \( T \) is \( S_2 \). Moreover,

\[
\dim(S_3 \cap S_2) = 1;
\]

therefore, \( T(v_0) = \lambda v_0 \) for some scalar \( \lambda \). Furthermore, let \( v_1 := T_1(v_0) \) and \( v_2 := T_2(v_1) = T(v_0) \). For vectors \( n_1 \) and \( n_2 \) in the orthogonal complement of \( S_1 \) and \( S_2 \), respectively, we have that \( n_1, n_2, v_0, v_1 \), and \( v_2 \) are coplanar, since they are in the 2-dimensional subspace \( S_3 \). Thus

\[
\langle v_0, v_1 \rangle = \langle v_1, v_2 \rangle = \langle n_1, n_2 \rangle = \theta.
\]

Hence, \( \cos \theta = \frac{\langle v_0, v_1 \rangle}{\|v_0\|\|v_1\|} \) and

\[
\|v_2\|\|v_1\| \cos \theta = \frac{\|v_2\|}{\|v_0\|} \langle v_0, v_1 \rangle = \lambda \langle v_0, v_1 \rangle.
\]

Note that \( \langle v_1, (v_0 - v_1) \rangle = \langle v_2, (v_1 - v_2) \rangle = 0 \), so

\[
\|v_2\|\|v_1\| \cos \theta = \lambda \langle v_1 + (v_0 - v_1), v_1 \rangle = \lambda \|v_1\|^2;
\]

thus \( \frac{\|v_2\|}{\|v_1\|} \cos \theta = \lambda \). Moreover,

\[
\|v_2\|\|v_1\| \cos \theta = \langle v_2, v_1 \rangle = \langle v_2, v_2 + (v_1 - v_2) \rangle = \|v_2\|^2,
\]

so \( \cos \theta = \frac{\|v_2\|}{\|v_1\|} \). It follows that \( \lambda = \cos^2 \theta \). \( \square \)

6. Conclusion

We have shown that for every finite-dimensional inner product space, the method of alternating orthogonal projections between two hyperplane subspaces \( S_1 \) and \( S_2 \) yields at most three distinct eigenvalues when we consider the composition of two orthogonal projections. Also, the eigenvectors of such a composition can be quickly identified to be in the subspaces \( S_1^\perp \), \( S_1 \cap S_2 \), and \( S_2 \cap (S_1 \cap S_2)^\perp \). We should mention the special, and somewhat trivial, cases where the angle between \( S_1 \) and \( S_2 \) is \( 0^\circ \) or \( 90^\circ \). In the case where \( \theta = 90^\circ \), we have that \( P_{S_2} P_{S_1} = P_{S_1 \cap S_2} \), and \( P_{S_2} P_{S_1} = P_{S_1} = P_{S_2} \) when \( \theta = 0^\circ \). In these cases, there are two distinct eigenvalues: 0 and 1. For \( \theta = 90^\circ \), the respective multiplicities are 2 and \( n - 2 \); for \( \theta = 0^\circ \), they are 1 and \( n - 1 \). It is also noteworthy that the multiplicities obtained in Theorem 5 guarantee that \( P_{S_2} P_{S_1} \) is nondefective, a necessary condition for the algorithm presented in Section 3.
Acknowledgements

This work was supported by a grant funded by the Division of Mathematical Sciences, National Science Foundation, award number 0502354. Special thanks go to Paul Barloon for his valuable support during this research project.

References


Received: 2012-05-31 Accepted: 2013-06-02

rholy2@gmail.com Johns Hopkins University, 29 Hillview Avenue, Boston, MA 02131, United States

marcolopez@my.unt.edu Department of Mathematics, University of North Texas, 1716 W. Hickory St. Apt 2, Denton, TX 76201, United States

douglas.mupasiri@uni.edu Department of Mathematics, University of Northern Iowa, 220 Wright Hall, Cedar Falls, IA 50614-0506, United States

mln40@msstate.edu Jackson State University, 8845 Hwy 12 West, Sallis, MS 39160, United States
The influence of education in reducing the HIV epidemic
Renee Margevicius and Hem Raj Joshi

127

On the zeros of $\zeta(s) - c$
Adam Boseman and Sebastian Pauli

137

Dynamic impact of a particle
Jeongho Ahn and Jared R. Wolf

147

Magic polygrams
Amanda Bienz, Karen A. Yokley and Crista Arangala

169

Trading cookies in a gambler’s ruin scenario
Kuejai Jungjaturapit, Timothy Pluta, Reza Rastegar, Alexander Roitershtein, Matthew Temba, Chad N. Vidden and Brian Wu

191

Decomposing induced characters of the centralizer of an $n$-cycle in the symmetric group on $2n$ elements
Joseph Ricci

221

On the geometric deformations of functions in $L^2[D]$
Luis Contreras, Derek Desantis and Kathryn Leonard

233

Spectral characterization for von Neumann’s iterative algorithm in $\mathbb{R}^n$
Rudy Joly, Marco López, Douglas Mupasiri and Michael Newsome

243

The 3-point Steiner problem on a cylinder
Denise M. Halverson and Andrew E. Logan

251