Appendix B

Simple matrices

Mathematicians also attempted to develop algebra of vectors but there was no natural definition of the product of two vectors that held in arbitrary dimensions. The first vector algebra that involved a noncommutative vector product (*that is,* v×w *need not equal* w×v) *was proposed by Hermann Grassmann in his book* Ausdehnungslehre (*1844*). *Grassmann's text also introduced the product of a column matrix and a row matrix, which resulted in what is now called a simple or a rank-one matrix. In the late 19th century the American mathematical physicist Willard Gibbs published his famous treatise on vector analysis. In that treatise Gibbs represented general matrices, which he called* dyadics, *as sums of simple matrices, which Gibbs called* dyads. *Later the physicist P. A. M. Dirac introduced the term "bra-ket" for what we now call the scalar product of a "bra"* (*row*) *vector times a "ket"* (*column*) *vector and the term "ket-bra" for the product of a ket times a bra, resulting in what we now call a simple matrix, as above. Our convention of identifying column matrices and vectors was introduced by physicists in the 20th century.*

−Suddhendu Biswas [52, p.2]

B.1 Rank-1 matrix (dyad)

Any matrix formed from the unsigned outer product of two vectors,

$$\Psi = uv^{\mathrm{T}} \in \mathbb{R}^{M\times N} \tag{1783}$$

where $u\in\mathbb{R}^M$ and $v\in\mathbb{R}^N$, is rank-1 and called *dyad.* Conversely, any rank-1 matrix must have the form Ψ . [233, prob.1.4.1] Product $-uv^{\mathrm{T}}$ is a *negative dyad.* For matrix products AB^{T}, in general, we have

$$\mathcal{R}(AB^{\mathrm{T}}) \subseteq \mathcal{R}(A)\,, \qquad \mathcal{N}(AB^{\mathrm{T}}) \supseteq \mathcal{N}(B^{\mathrm{T}}) \tag{1784}$$

with equality when $B=A$ [374, §3.3, §3.6][B.1] or respectively when B is invertible and $\mathcal{N}(A)=\mathbf{0}$. Yet for all nonzero dyads we have

$$\mathcal{R}(uv^{\mathrm{T}}) = \mathcal{R}(u)\,, \qquad \mathcal{N}(uv^{\mathrm{T}}) = \mathcal{N}(v^{\mathrm{T}}) \equiv v^{\perp} \tag{1785}$$

[B.1] **Proof.** $\mathcal{R}(AA^{\mathrm{T}}) \subseteq \mathcal{R}(A)$ is obvious.

$$\begin{aligned}\mathcal{R}(AA^{\mathrm{T}}) &= \{AA^{\mathrm{T}}y \mid y\in\mathbb{R}^m\}\\ &\supseteq \{AA^{\mathrm{T}}y \mid A^{\mathrm{T}}y\in\mathcal{R}(A^{\mathrm{T}})\} = \mathcal{R}(A) \text{ by (146)}\end{aligned} \quad \blacklozenge$$

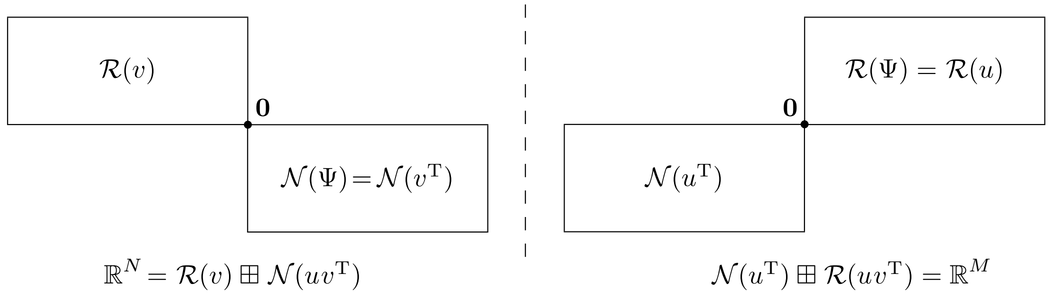


Figure 183: The four fundamental subspaces [376, §3.6] for any dyad $\Psi = uv^{\mathrm{T}} \in \mathbb{R}^{M\times N}$: $\mathcal{R}(v) \perp \mathcal{N}(\Psi)$ & $\mathcal{N}(u^{\mathrm{T}}) \perp \mathcal{R}(\Psi)$. $\Psi(x) \triangleq uv^{\mathrm{T}}x$ is a linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$. Map from $\mathcal{R}(v)$ to $\mathcal{R}(u)$ is bijective. [374, §3.1]

where $\dim v^{\perp} = N-1$.

It is obvious that a dyad can be $\mathbf{0}$ only when u or v is $\mathbf{0}$;

$$\Psi = uv^{\mathrm{T}} = \mathbf{0} \quad \Leftrightarrow \quad u = \mathbf{0} \;\; \textbf{or} \;\; v = \mathbf{0} \tag{1786}$$

The matrix 2-norm for Ψ is equivalent to Frobenius' norm;

$$\|\Psi\|_2 = \sigma_1 = \|uv^{\mathrm{T}}\|_{\mathrm{F}} = \|uv^{\mathrm{T}}\|_2 = \|u\|\,\|v\| \tag{1787}$$

When u and v are normalized, the pseudoinverse is the transposed dyad. Otherwise,

$$\Psi^{\dagger} = (uv^{\mathrm{T}})^{\dagger} = \frac{vu^{\mathrm{T}}}{\|u\|^2\,\|v\|^2} \tag{1788}$$

When dyad $uv^{\mathrm{T}} \in \mathbb{R}^{N\times N}$ is square, uv^{T} has at least $N-1$ 0-eigenvalues and corresponding eigenvectors spanning $v^{\perp}$. The remaining eigenvector u spans the range of uv^{T} with corresponding eigenvalue

$$\lambda = v^{\mathrm{T}}u = \mathrm{tr}(uv^{\mathrm{T}}) \in \mathbb{R} \tag{1789}$$

Determinant is a product of the eigenvalues; so, it is always true that

$$\det \Psi = \det(uv^{\mathrm{T}}) = 0 \tag{1790}$$

When $\lambda = 1$, the square dyad is a nonorthogonal projector projecting on its range ($\Psi^2 = \Psi$, §E.6); a *projector dyad.* It is quite possible that $u \in v^{\perp}$ making the remaining eigenvalue instead 0;[B.2] $\lambda = 0$ together with the first $N-1$ 0-eigenvalues; *id est*, it is possible uv^{T} were nonzero while all its eigenvalues are 0. The matrix

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \tag{1791}$$

for example, has two 0-eigenvalues. In other words, eigenvector u may simultaneously be a member of the nullspace and range of the dyad. The explanation is, simply, because u and v share the same dimension, $\dim u = M = \dim v = N$:

[B.2] A dyad is not always diagonalizable (§A.5) because its eigenvectors are not necessarily independent.

Proof. Figure **183** sees the four fundamental subspaces for the dyad. Linear operator $\Psi : \mathbb{R}^N \rightarrow \mathbb{R}^M$ provides a map between vector spaces that remain distinct when $M = N$;

$$\begin{aligned} u &\in \mathcal{R}(uv^{\mathrm{T}}) \\ u \in \mathcal{N}(uv^{\mathrm{T}}) &\Leftrightarrow v^{\mathrm{T}}u = 0 \\ \mathcal{R}(uv^{\mathrm{T}}) \cap \mathcal{N}(uv^{\mathrm{T}}) &= \emptyset \end{aligned} \tag{1792}$$

♦

B.1.0.1 rank-1 modification

For $A \in \mathbb{R}^{M\times N}$, $x \in \mathbb{R}^N$, $y \in \mathbb{R}^M$, and $y^{\mathrm{T}}Ax \neq 0$ [238, §2.1][B.3]

$$\operatorname{rank}\left(A - \frac{Axy^{\mathrm{T}}A}{y^{\mathrm{T}}Ax}\right) = \operatorname{rank}(A) - 1 \tag{1794}$$

Given nonsingular matrix $A \in \mathbb{R}^{N\times N} \ni 1 + v^{\mathrm{T}}A^{-1}u \neq 0$, [178, §4.11.2] [253, App.6] [463, §2.3 prob.16] (Sherman-Morrison-Woodbury)

$$(A \pm uv^{\mathrm{T}})^{-1} = A^{-1} \mp \frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1 \pm v^{\mathrm{T}}A^{-1}u} \tag{1795}$$

B.1.0.2 dyad symmetry

In the specific circumstance that $v = u$, then $uu^{\mathrm{T}} \in \mathbb{R}^{N\times N}$ is symmetric, rank-1 , and positive semidefinite having exactly $N-1$ 0-eigenvalues. In fact, (Theorem A.3.1.0.7)

$$uv^{\mathrm{T}} \succeq 0 \;\Leftrightarrow\; v = u \tag{1796}$$

and the remaining eigenvalue is almost always positive;

$$\lambda \;=\; u^{\mathrm{T}}u \;=\; \operatorname{tr}(uu^{\mathrm{T}}) > 0 \;\text{ unless }\; u = \mathbf{0} \tag{1797}$$

When $\lambda = 1$, the dyad becomes an orthogonal projector.

Matrix

$$\begin{bmatrix} \Psi & u \\ u^{\mathrm{T}} & 1 \end{bmatrix} \tag{1798}$$

for example, is rank-1 positive semidefinite if and only if $\Psi = uu^{\mathrm{T}}$.

B.1.1 Dyad independence

Now we consider a sum of dyads like (1783) as encountered in diagonalization and singular value decomposition:

$$\mathcal{R}\left(\sum_{i=1}^{k} s_i w_i^{\mathrm{T}}\right) = \sum_{i=1}^{k} \mathcal{R}\big(s_i w_i^{\mathrm{T}}\big) = \sum_{i=1}^{k} \mathcal{R}(s_i) \;\Leftarrow\; w_i \;\forall\, i \;\text{ are l.i.} \tag{1799}$$

[B.3] This rank-1 modification formula has a Schur progenitor, in the symmetric case:

$$\begin{array}{ll} \underset{c}{\text{minimize}} & c \\ \text{subject to} & \begin{bmatrix} A & Ax \\ y^{\mathrm{T}}A & c \end{bmatrix} \succeq 0 \end{array} \tag{1793}$$

has analytical solution by (1679b): $c \geq y^{\mathrm{T}}AA^{\dagger}Ax = y^{\mathrm{T}}Ax$. Difference $A - \frac{Axy^{\mathrm{T}}A}{y^{\mathrm{T}}Ax}$ comes from (1679c). Rank modification is provable via Theorem A.4.0.1.3.

range of summation is the vector sum of ranges.[B.4] (Theorem B.1.1.1.1) Under the assumption the dyads are linearly independent (l.i.), then vector sums are unique (p.631): for $\{w_i\}$ l.i. and $\{s_i\}$ l.i.

$$\mathcal{R}\left(\sum_{i=1}^{k} s_i w_i^{\mathrm T}\right) = \mathcal{R}\big(s_1 w_1^{\mathrm T}\big) \oplus \ldots \oplus \mathcal{R}\big(s_k w_k^{\mathrm T}\big) \;=\; \mathcal{R}(s_1) \oplus \ldots \oplus \mathcal{R}(s_k) \tag{1800}$$

B.1.1.0.1 Definition. *Linearly independent dyads.* [243, p.29 thm.11] [382, p.2]
The set of k dyads

$$\left\{ s_i w_i^{\mathrm T} \mid i=1\ldots k \right\} \tag{1801}$$

where $s_i \in \mathbb{C}^M$ and $w_i \in \mathbb{C}^N$, is said to be linearly independent iff

$$\operatorname{rank}\left(SW^{\mathrm T} \triangleq \sum_{i=1}^{k} s_i w_i^{\mathrm T} \right) \;=\; k \tag{1802}$$

where $S \triangleq [s_1 \cdots s_k] \in \mathbb{C}^{M\times k}$ and $W \triangleq [w_1 \cdots w_k] \in \mathbb{C}^{N\times k}$. △

Dyad independence does not preclude existence of a nullspace $\mathcal{N}(SW^{\mathrm T})$, as defined, nor does it imply $SW^{\mathrm T}$ were full-rank. In absence of assumption of independence, generally, $\operatorname{rank} SW^{\mathrm T} \le k$. Conversely, any rank-$k$ matrix can be written in the form $SW^{\mathrm T}$ by singular value decomposition. (§A.6)

B.1.1.0.2 Theorem. *Linearly independent (l.i.) dyads.*
Vectors $\{s_i \in \mathbb{C}^M,\ i=1\ldots k\}$ are l.i. and vectors $\{w_i \in \mathbb{C}^N,\ i=1\ldots k\}$ are l.i. if and only if dyads $\{s_i w_i^{\mathrm T} \in \mathbb{C}^{M\times N},\ i=1\ldots k\}$ are l.i. ⋄

Proof. Linear independence of k dyads is identical to definition (1802).
(⇒) Suppose $\{s_i\}$ and $\{w_i\}$ are each linearly independent sets. Invoking Sylvester's rank inequality, [233, §0.4] [463, §2.4]

$$\operatorname{rank} S + \operatorname{rank} W \,-\, k \;\le\; \operatorname{rank}(SW^{\mathrm T}) \;\le\; \min\{\operatorname{rank} S\,,\, \operatorname{rank} W\} \;\;(\le\; k) \tag{1803}$$

Then $k \le \operatorname{rank}(SW^{\mathrm T}) \le k$ which implies the dyads are independent.
(⇐) Conversely, suppose $\operatorname{rank}(SW^{\mathrm T}) = k$. Then

$$k \le \min\{\operatorname{rank} S\,,\, \operatorname{rank} W\} \le k \tag{1804}$$

implying the vector sets are each independent. ♦

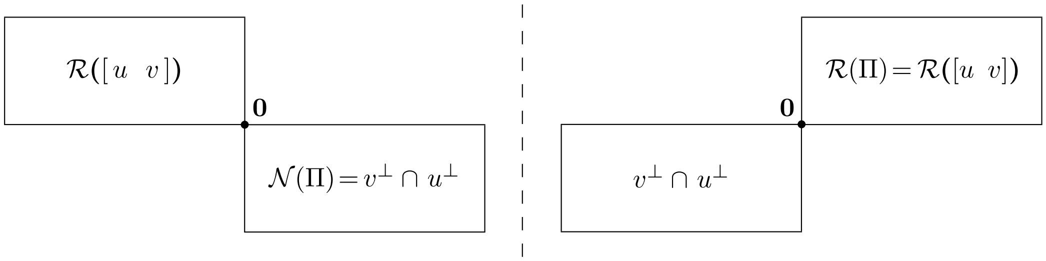
B.1.1.1 Biorthogonality condition, Range and Nullspace of Sum

Dyads characterized by biorthogonality condition $W^{\mathrm T}S = I$ are independent; *id est*, for $S \in \mathbb{C}^{M\times k}$ and $W \in \mathbb{C}^{N\times k}$, if $W^{\mathrm T}S = I$ then $\operatorname{rank}(SW^{\mathrm T}) = k$ by the *linearly independent dyads theorem* because (*confer* §E.1.1)

$$W^{\mathrm T}S = I \;\;\Rightarrow\;\; \operatorname{rank} S = \operatorname{rank} W = k \le M = N \tag{1805}$$

To see that, we need only show: $\mathcal{N}(S) = \mathbf{0} \;\Leftrightarrow\; \exists\, B \;\ni\; BS = I$.[B.5]
(⇐) Assume $BS = I$. Then $\mathcal{N}(BS) = \mathbf{0} = \{x \mid BSx = \mathbf{0}\} \supseteq \mathcal{N}(S)$. (1784)

[B.4] Move of range $\mathcal{R}$ to inside summation admitted by linear independence of $\{w_i\}$.
[B.5] Left inverse is not unique, in general.



$$\mathbb{R}^N=\mathcal{R}([u\;\;v])\boxplus\mathcal{N}\left(\begin{bmatrix}v^{\mathrm T}\\ u^{\mathrm T}\end{bmatrix}\right)\qquad\qquad \mathcal{N}\left(\begin{bmatrix}v^{\mathrm T}\\ u^{\mathrm T}\end{bmatrix}\right)\boxplus\mathcal{R}([u\;\;v])=\mathbb{R}^N$$

Figure 184: The four fundamental subspaces [376, §3.6] for doublet $\Pi=uv^{\mathrm T}+vu^{\mathrm T}\in\mathbb{S}^N$. $\Pi(x)=(uv^{\mathrm T}+vu^{\mathrm T})x$ is a linear bijective mapping from $\mathcal{R}([u\;\;v])$ to $\mathcal{R}([u\;\;v])$.

(⇒) If $\mathcal{N}(S)=\mathbf{0}$ then S must be full-rank thin-or-square.
$\therefore\;\exists\,A\,,B\,,C\;\ni\;\begin{bmatrix}B\\ C\end{bmatrix}[\,S\;\;A\,]=I$ (*id est*, $[\,S\;\;A\,]$ is invertible) $\Rightarrow\;BS=I$.
Left inverse B is given as $W^{\mathrm T}$ here. Because of reciprocity with S, it immediately follows: $\mathcal{N}(W)=\mathbf{0}\;\Leftrightarrow\;\exists\,S\;\ni\;S^{\mathrm T}W=I$. ♦

Dyads produced by diagonalization, for example, are independent because of their inherent biorthogonality. (§A.5.0.3) The converse is generally false; *id est*, linearly independent dyads are not necessarily biorthogonal.

B.1.1.1.1 Theorem. *Nullspace and range of dyad sum.*
Given a sum of dyads represented by $SW^{\mathrm T}$ where $S\in\mathbb{C}^{M\times k}$ and $W\in\mathbb{C}^{N\times k}$

$$\begin{aligned}\mathcal{N}(SW^{\mathrm T})&=\mathcal{N}(W^{\mathrm T})\quad\Leftarrow\quad\exists\,B\;\ni\;BS=I\\ \mathcal{R}(SW^{\mathrm T})&=\mathcal{R}(S)\quad\Leftarrow\quad\exists\,Z\;\ni\;W^{\mathrm T}Z=I\end{aligned}\tag{1806}$$

⋄

Proof. (⇒) $\mathcal{N}(SW^{\mathrm T})\supseteq\mathcal{N}(W^{\mathrm T})$ and $\mathcal{R}(SW^{\mathrm T})\subseteq\mathcal{R}(S)$ are obvious.
(⇐) Assume existence of a left inverse $B\in\mathbb{R}^{k\times N}$ and a right inverse $Z\in\mathbb{R}^{N\times k}$.[B.6]

$$\mathcal{N}(SW^{\mathrm T})\;=\;\{x\mid SW^{\mathrm T}x=\mathbf{0}\}\;\subseteq\;\{x\mid BSW^{\mathrm T}x=\mathbf{0}\}\;=\;\mathcal{N}(W^{\mathrm T})\tag{1807}$$
$$\mathcal{R}(SW^{\mathrm T})\;=\;\{SW^{\mathrm T}x\mid x\in\mathbb{R}^N\}\;\supseteq\;\{SW^{\mathrm T}Zy\mid Zy\in\mathbb{R}^N\}\;=\;\mathcal{R}(S)\tag{1808}$$

♦

B.2 Doublet

Consider a sum of two linearly independent square dyads, one a transposition of the other:

$$\Pi\;=\;uv^{\mathrm T}+vu^{\mathrm T}\;=\;[\,u\;\;v\,]\begin{bmatrix}v^{\mathrm T}\\ u^{\mathrm T}\end{bmatrix}\;=\;SW^{\mathrm T}\;\in\;\mathbb{S}^N\tag{1809}$$

where $u\,,v\in\mathbb{R}^N$. Like the dyad, a doublet can be $\mathbf{0}$ only when u or v is $\mathbf{0}$;

$$\Pi=uv^{\mathrm T}+vu^{\mathrm T}=\mathbf{0}\;\Leftrightarrow\;u=\mathbf{0}\;\;\textbf{or}\;\;v=\mathbf{0}\tag{1810}$$

By assumption of independence, a nonzero doublet has two nonzero eigenvalues

$$\lambda_1\triangleq u^{\mathrm T}v+\|uv^{\mathrm T}\|\;,\qquad\lambda_2\triangleq u^{\mathrm T}v-\|uv^{\mathrm T}\|\tag{1811}$$

[B.6] By counterexample, the theorem's converse cannot be true; *e.g*, $S=W=[\,\mathbf{1}\;\;\mathbf{0}\,]$.

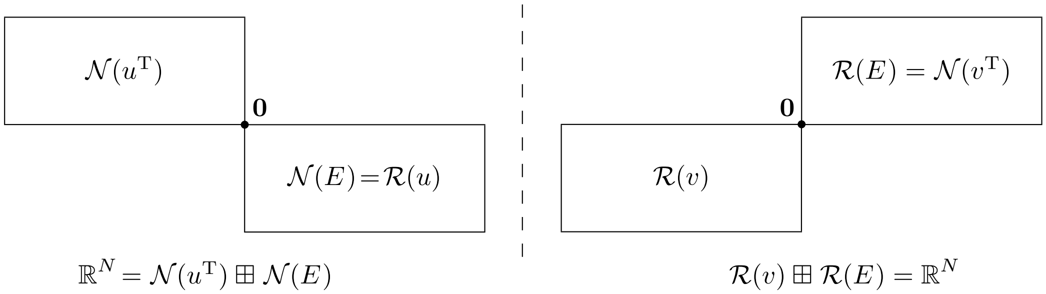


Figure 185: $v^{\mathrm T}u=1/\zeta$. The four fundamental subspaces [376, §3.6] for elementary matrix E as a linear mapping $E(x)=\left(I-\dfrac{uv^{\mathrm T}}{v^{\mathrm T}u}\right)x$.

where $\lambda_1>0>\lambda_2$, with corresponding eigenvectors

$$x_1 \triangleq \frac{u}{\|u\|}+\frac{v}{\|v\|}\ , \qquad x_2 \triangleq \frac{u}{\|u\|}-\frac{v}{\|v\|} \tag{1812}$$

spanning the doublet range. Eigenvalue λ_1 cannot be 0 unless u and v have opposing directions, but that is antithetical since then the dyads would no longer be independent. Eigenvalue λ_2 is 0 if and only if u and v share the same direction, again antithetical. Generally we have $\lambda_1>0$ and $\lambda_2<0$, so Π is indefinite.

By the *nullspace and range of dyad sum theorem,* doublet Π has $N-2$ zero-eigenvalues remaining and corresponding eigenvectors spanning $\mathcal{N}\left(\begin{bmatrix} v^{\mathrm T}\\ u^{\mathrm T}\end{bmatrix}\right)$. We therefore have

$$\mathcal{R}(\Pi)=\mathcal{R}([\,u\ \ v\,])\ , \qquad \mathcal{N}(\Pi)=v^{\perp}\cap u^{\perp} \tag{1813}$$

of respective dimension 2 and $N-2$. (Figure **184**)

B.3 Elementary matrix

A matrix of the form

$$E=I-\zeta\, uv^{\mathrm T}\in\mathbb{R}^{N\times N} \tag{1814}$$

where $\zeta\in\mathbb{R}$ is finite and $u,v\in\mathbb{R}^N$, is called *elementary matrix* or *rank*-1 *modification of the Identity*. [235] Any elementary matrix in $\mathbb{R}^{N\times N}$ has $N-1$ eigenvalues equal to 1 corresponding to real eigenvectors that span $v^{\perp}$. The remaining eigenvalue

$$\lambda=1-\zeta\, v^{\mathrm T}u \tag{1815}$$

corresponds to eigenvector u .[B.7] From [253, App.7.A.26] the determinant:

$$\det E \;=\; 1-\operatorname{tr}\!\left(\zeta\, uv^{\mathrm T}\right) \;=\; \lambda \tag{1816}$$

If $\lambda\neq 0$ then E is invertible; [178] (*confer* §B.1.0.1)

$$E^{-1}=I+\frac{\zeta}{\lambda}uv^{\mathrm T} \tag{1817}$$

Eigenvectors corresponding to 0 eigenvalues belong to $\mathcal{N}(E)$, and the number of 0 eigenvalues must be at least $\dim\mathcal{N}(E)$ which, here, can be at most one.

[B.7] Elementary matrix E is not always diagonalizable because eigenvector u need not be independent of the others; *id est,* $u\in v^{\perp}$ is possible.

(§A.7.3.0.1) The nullspace exists, therefore, when $\lambda=0$; *id est*, when $v^{\mathrm T}u=1/\zeta$; rather, whenever u belongs to hyperplane $\{z\in\mathbb{R}^N \mid v^{\mathrm T}z=1/\zeta\}$. Then (when $\lambda=0$) elementary matrix E is a nonorthogonal projector projecting on its range ($E^2=E$, §E.1) and $\mathcal{N}(E)=\mathcal{R}(u)$; eigenvector u spans the nullspace when it exists. By conservation of dimension, $\dim\mathcal{R}(E)=N-\dim\mathcal{N}(E)$. It is apparent from (1814) that $v^\perp\subseteq\mathcal{R}(E)$, but $\dim v^\perp=N-1$. Hence $\mathcal{R}(E)\equiv v^\perp$ when the nullspace exists, and the remaining eigenvectors span it.

In summary, when a nontrivial nullspace of E exists,

$$\mathcal{R}(E)=\mathcal{N}(v^{\mathrm T})\,,\qquad \mathcal{N}(E)=\mathcal{R}(u)\,,\qquad v^{\mathrm T}u=1/\zeta \tag{1818}$$

illustrated in Figure **185**, which is opposite to the assignment of subspaces for a dyad (Figure **183**). Otherwise, $\mathcal{R}(E)=\mathbb{R}^N$.

When $E=E^{\mathrm T}$, the spectral norm is

$$\|E\|_2=\max\{1\,,\,|\lambda|\} \tag{1819}$$

B.3.1 Householder matrix

An elementary matrix is called a Householder matrix when it has the defining form, for nonzero vector u [185, §5.1.2] [178, §4.10.1] [374, §7.3] [233, §2.2]

$$H=I-2\frac{uu^{\mathrm T}}{u^{\mathrm T}u}\in\mathbb{S}^N \tag{1820}$$

which is a symmetric orthogonal (reflection) matrix $\big(H^{-1}=H^{\mathrm T}=H$ (§B.5.3)$\big)$. Vector u is normal to an $N-1$-dimensional subspace $u^\perp$ through which this particular H effects pointwise reflection; *e.g*, $Hu^\perp=u^\perp$ while $Hu=-u$.

Matrix H has $N-1$ orthonormal eigenvectors spanning that reflecting subspace $u^\perp$ with corresponding eigenvalues equal to 1. The remaining eigenvector u has corresponding eigenvalue -1; so

$$\det H=-1 \tag{1821}$$

Due to symmetry of H, the matrix 2-norm (the spectral norm) is equal to the largest eigenvalue-magnitude. A Householder matrix is thus characterized,

$$H^{\mathrm T}=H\,,\qquad H^{-1}=H^{\mathrm T}\,,\qquad \|H\|_2=1\,,\qquad H\nsucceq 0 \tag{1822}$$

For example, the permutation matrix

$$\Xi=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix} \tag{1823}$$

is a Householder matrix having $u=[\,0\;\;1\;\;-1\,]^{\mathrm T}/\sqrt{2}$. Not all permutation matrices are Householder matrices, although all permutation matrices are orthogonal matrices $\big($§B.5.2, $\Xi^{\mathrm T}\Xi=I\big)$ [374, §3.4] because they are made by permuting rows and columns of the Identity matrix. Neither are all symmetric permutation matrices Householder matrices; *e.g*, $\Xi=\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}$ (1920) is not a Householder matrix.

B.4 Auxiliary V-matrices

B.4.1 Auxiliary projector matrix V

It is convenient to define a matrix V that arises naturally as a consequence of translating geometric center α_c (§5.5.1.0.1) of some list X to the origin. In place of $X-\alpha_c\mathbf{1}^{\mathrm T}$ we may write XV as in (1135) where

$$V \;=\; I-\frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm T}\in\mathbb{S}^N \qquad (1071)$$

is an elementary matrix called the *geometric centering matrix.*

Any elementary matrix in $\mathbb{R}^{N\times N}$ has $N-1$ eigenvalues equal to 1. For the particular elementary matrix V, the N^{th} eigenvalue equals 0. The number of 0 eigenvalues must equal $\dim\mathcal{N}(V)=1$, by the 0 *eigenvalues theorem* (§A.7.3.0.1), because $V=V^{\mathrm T}$ is diagonalizable. Because

$$V\mathbf{1}=\mathbf{0} \qquad (1824)$$

the nullspace $\mathcal{N}(V)=\mathcal{R}(\mathbf{1})$ is spanned by the eigenvector $\mathbf{1}$. The remaining eigenvectors span $\mathcal{R}(V)\equiv\mathbf{1}^{\perp}=\mathcal{N}(\mathbf{1}^{\mathrm T})$ that has dimension $N-1$.

Because

$$V^2=V \qquad (1825)$$

and $V^{\mathrm T}=V$, elementary matrix V is also a projection matrix (§E.3) projecting orthogonally on its range $\mathcal{N}(\mathbf{1}^{\mathrm T})$ which is a hyperplane containing the origin in $\mathbb{R}^N$

$$V=I-\mathbf{1}(\mathbf{1}^{\mathrm T}\mathbf{1})^{-1}\mathbf{1}^{\mathrm T} \qquad (1826)$$

The $\{0,1\}$ eigenvalues also indicate that diagonalizable V is a projection matrix. [463, §4.1 thm.4.1] Symmetry of V denotes orthogonal projection; from (2120),

$$V^2=V\;,\qquad V^{\mathrm T}=V\;,\qquad V^{\dagger}=V\;,\qquad \|V\|_2=1\,,\qquad V\succeq 0 \qquad (1827)$$

Matrix V is also circulant [197].

B.4.1.0.1 Example. *Relationship of Auxiliary to Householder matrix.*
Let $H\in\mathbb{S}^N$ be a Householder matrix (1820) defined by

$$u=\begin{bmatrix}1\\ \vdots\\ 1\\ 1+\sqrt{N}\end{bmatrix}\in\mathbb{R}^N \qquad (1828)$$

Then we have [180, §2]

$$V=H\begin{bmatrix}I & \mathbf{0}\\ \mathbf{0}^{\mathrm T} & 0\end{bmatrix}H \qquad (1829)$$

Let $D\in\mathbb{S}_h^N$ and define

$$-HDH\triangleq-\begin{bmatrix}A & b\\ b^{\mathrm T} & c\end{bmatrix} \qquad (1830)$$

where b is a vector. Then because H is nonsingular (§A.3.1.0.5) [215, §3]

$$-VDV=-H\begin{bmatrix}A & \mathbf{0}\\ \mathbf{0}^{\mathrm T} & 0\end{bmatrix}H\succeq 0\;\Leftrightarrow\;-A\succeq 0 \qquad (1831)$$

and affine dimension is $r=\operatorname{rank}A$ when D is a Euclidean distance matrix. □

B.4.2 Schoenberg auxiliary matrix $V_{\mathcal{N}}$

1. $V_{\mathcal{N}} = \frac{1}{\sqrt{2}}\begin{bmatrix} -\mathbf{1}^{\mathrm T} \\ I \end{bmatrix} \in \mathbb{R}^{N\times N-1}$ (1055)

2. $V_{\mathcal{N}}^{\mathrm T}\mathbf{1} = \mathbf{0}$

3. $I - e_1\mathbf{1}^{\mathrm T} = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$

4. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} V_{\mathcal{N}} = V_{\mathcal{N}}$

5. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} V = V$

6. $V\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$

7. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$

8. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^{\dagger} = \begin{bmatrix} 0 & \mathbf{0}^{\mathrm T} \\ \mathbf{0} & I \end{bmatrix} V$

9. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^{\dagger} V = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^{\dagger}$

10. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^{\dagger} = V$

11. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^{\dagger}\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}^{\mathrm T} \\ \mathbf{0} & I \end{bmatrix}$

12. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}\begin{bmatrix} 0 & \mathbf{0}^{\mathrm T} \\ \mathbf{0} & I \end{bmatrix} = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$

13. $\begin{bmatrix} 0 & \mathbf{0}^{\mathrm T} \\ \mathbf{0} & I \end{bmatrix}\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}^{\mathrm T} \\ \mathbf{0} & I \end{bmatrix}$

14. $[\,V_{\mathcal{N}} \quad \frac{1}{\sqrt{2}}\mathbf{1}\,]^{-1} = \begin{bmatrix} V_{\mathcal{N}}^{\dagger} \\ \frac{\sqrt{2}}{N}\mathbf{1}^{\mathrm T} \end{bmatrix}$

15. $V_{\mathcal{N}}^{\dagger} = \sqrt{2}\begin{bmatrix}-\frac{1}{N}\mathbf{1} & I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm T}\end{bmatrix} \in \mathbb{R}^{N-1\times N}$, $\left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm T} \in \mathbb{S}^{N-1}\right)$

16. $V_{\mathcal{N}}^{\dagger}\mathbf{1} = \mathbf{0}$

17. $V_{\mathcal{N}}^{\dagger}V_{\mathcal{N}} = I\,,\;\; V_{\mathcal{N}}^{\mathrm T}V_{\mathcal{N}} = \frac{1}{2}(I + \mathbf{1}\mathbf{1}^{\mathrm T}) \in \mathbb{S}^{N-1}$

18. $V^{\mathrm T} = V = V_{\mathcal{N}}V_{\mathcal{N}}^{\dagger} = I - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm T} \in \mathbb{S}^N$

19. $-V_{\mathcal{N}}^{\dagger}(\mathbf{1}\mathbf{1}^{\mathrm T} - I)V_{\mathcal{N}} = I$, $\left(\mathbf{1}\mathbf{1}^{\mathrm T} - I \in \mathbb{EDM}^N\right)$

20. $D = [d_{ij}] \in \mathbb{S}_h^N$ (1073)

    $$\operatorname{tr}(-VDV) = \operatorname{tr}(-VD) = \operatorname{tr}(-V_{\mathcal{N}}^{\dagger}DV_{\mathcal{N}}) = \tfrac{1}{N}\mathbf{1}^{\mathrm T}D\,\mathbf{1} = \tfrac{1}{N}\operatorname{tr}(\mathbf{1}\mathbf{1}^{\mathrm T}D) = \tfrac{1}{N}\sum_{i,j} d_{ij}$$

    Any elementary matrix $E \in \mathbb{S}^N$ of the particular form

    $$E = k_1 I - k_2\,\mathbf{1}\mathbf{1}^{\mathrm T} \tag{1832}$$

    where $k_1\,, k_2 \in \mathbb{R}$,[B.8] will make $\operatorname{tr}(-ED)$ proportional to $\sum d_{ij}$.

[B.8] If k_1 is $1-\rho$ while k_2 equals $-\rho\in\mathbb{R}$, then all eigenvalues of E for $-1/(N-1) < \rho < 1$ are guaranteed positive and therefore E is guaranteed positive definite. [343]

21. $D=[d_{ij}]\in\mathbb{S}^N$

$$\operatorname{tr}(-VDV) \;=\; \tfrac{1}{N}\sum_{\substack{i,j\\ i\neq j}} d_{ij} - \tfrac{N-1}{N}\sum_i d_{ii} \;=\; \tfrac{1}{N}\mathbf{1}^{\rm T}D\,\mathbf{1} - \operatorname{tr}D$$

22. $D=[d_{ij}]\in\mathbb{S}_h^N$

$$\operatorname{tr}(-V_{\mathcal{N}}^{\rm T}DV_{\mathcal{N}}) = \sum_j d_{1j}$$

23. For $Y\in\mathbb{S}^N$

$$V\big(Y-\delta(Y\mathbf{1})\big)V = Y-\delta(Y\mathbf{1})$$

B.4.3 Orthonormal auxiliary matrix $V_{\mathcal{W}}$

Thin matrix

$$V_{\mathcal{W}} \triangleq \begin{bmatrix} \frac{-1}{\sqrt{N}} & \frac{-1}{\sqrt{N}} & \cdots & \frac{-1}{\sqrt{N}} \\ 1+\frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & \frac{-1}{N+\sqrt{N}} \\ \frac{-1}{N+\sqrt{N}} & \ddots & \ddots & \frac{-1}{N+\sqrt{N}} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & 1+\frac{-1}{N+\sqrt{N}} \end{bmatrix} \in \mathbb{R}^{N\times N-1} \tag{1833}$$

has $\mathcal{R}(V_{\mathcal{W}})=\mathcal{N}(\mathbf{1}^{\rm T})$ and orthonormal columns. [7] We defined three auxiliary V-matrices: V, $V_{\mathcal{N}}$ (1055), and $V_{\mathcal{W}}$ sharing some attributes listed in Table **B.4.4**. For example, V can be expressed

$$V = V_{\mathcal{W}}V_{\mathcal{W}}^{\rm T} = V_{\mathcal{N}}V_{\mathcal{N}}^{\dagger} \tag{1834}$$

but $V_{\mathcal{W}}^{\rm T}V_{\mathcal{W}} = I$ means V is an orthogonal projector (2117) and

$$V_{\mathcal{W}}^{\dagger} = V_{\mathcal{W}}^{\rm T}\,, \qquad \|V_{\mathcal{W}}\|_2 = 1\,, \qquad V_{\mathcal{W}}^{\rm T}\mathbf{1} = \mathbf{0} \tag{1835}$$

B.4.4 Auxiliary V-matrix Table

	$\dim V$	$\operatorname{rank} V$	$\mathcal{R}(V)$	$\mathcal{N}(V^{\rm T})$	$V^{\rm T}V$	$VV^{\rm T}$	$VV^{\dagger}$
V	$N\times N$	$N-1$	$\mathcal{N}(\mathbf{1}^{\rm T})$	$\mathcal{R}(\mathbf{1})$	V	V	V
$V_{\mathcal{N}}$	$N\times(N-1)$	$N-1$	$\mathcal{N}(\mathbf{1}^{\rm T})$	$\mathcal{R}(\mathbf{1})$	$\frac{1}{2}(I+\mathbf{1}\mathbf{1}^{\rm T})$	$\frac{1}{2}\begin{bmatrix} N-1 & -\mathbf{1}^{\rm T} \\ -\mathbf{1} & I \end{bmatrix}$	V
$V_{\mathcal{W}}$	$N\times(N-1)$	$N-1$	$\mathcal{N}(\mathbf{1}^{\rm T})$	$\mathcal{R}(\mathbf{1})$	I	V	V

B.4.5 More auxiliary matrices

Mathar shows [295, §2] that any elementary matrix (§B.3) of the form

$$V_{\mathcal{S}} = I - b\,\mathbf{1}^{\rm T} \in \mathbb{R}^{N\times N} \tag{1836}$$

such that $b^{\rm T}\mathbf{1}=1$ (*confer* [190, §2]), is an auxiliary V-matrix having

$$\begin{array}{rr} \mathcal{R}(V_{\mathcal{S}}^{\rm T}) = \mathcal{N}(b^{\rm T})\,, & \mathcal{R}(V_{\mathcal{S}}) = \mathcal{N}(\mathbf{1}^{\rm T}) \\ \mathcal{N}(V_{\mathcal{S}}) = \mathcal{R}(b)\,, & \mathcal{N}(V_{\mathcal{S}}^{\rm T}) = \mathcal{R}(\mathbf{1}) \end{array} \tag{1837}$$

Given $X \in \mathbb{R}^{n\times N}$, the choice $b = \frac{1}{N}\mathbf{1}$ $(V_{\mathcal{S}} = V)$ minimizes $\|X(I - b\,\mathbf{1}^{\mathrm T})\|_{\mathrm F}$. [192, §3.2.1]

B.5 Orthomatrices

B.5.1 Orthonormal matrix

Property $Q^{\mathrm T}Q = I$ completely defines orthonormal matrix $Q \in \mathbb{R}^{n\times k}$ $(k \leq n)$; a full-rank thin-or-square matrix characterized by nonexpansivity (2121)

$$\|Q^{\mathrm T}x\|_2 \leq \|x\|_2 \quad \forall\, x \in \mathbb{R}^n, \qquad \|Qy\|_2 = \|y\|_2 \quad \forall\, y \in \mathbb{R}^k \tag{1838}$$

and preservation of vector inner-product

$$\langle Qy\,,\, Qz\rangle = \langle y\,,\, z\rangle \tag{1839}$$

B.5.2 Orthogonal matrix & vector rotation

An orthogonal matrix is a square orthonormal matrix. Property $Q^{-1} = Q^{\mathrm T}$ completely defines orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ employed to effect vector rotation; [374, §2.6, §3.4] [376, §6.5] [233, §2.1] for any $x \in \mathbb{R}^n$

$$\|Qx\|_2 = \|x\|_2 \tag{1840}$$

In other words, the 2-norm is orthogonally invariant. Any antisymmetric matrix constructs an orthogonal matrix; *id est*, for $A = -A^{\mathrm T}$

$$Q = (I + A)^{-1}(I - A) \tag{1841}$$

A *unitary matrix* is a complex generalization of orthogonal matrix; conjugate transpose defines it: $U^{-1} = U^{\mathrm H}$. An orthogonal matrix is simply a real unitary matrix.[B.9]

Orthogonal matrix Q is a normal matrix further characterized by norm:

$$Q^{-1} = Q^{\mathrm T}\,, \qquad \|Q\|_2^2 = 1\,, \qquad \|Q\|_{\mathrm F}^2 = n \tag{1842}$$

Applying this characterization to $Q^{\mathrm T}$, we see that it too is an orthogonal matrix because transpose inverse equals inverse transpose. Hence the rows and columns of Q each respectively form an orthonormal set. Normalcy guarantees diagonalization (§A.5.1.0.1). So, for $Q \triangleq S\Lambda S^{\mathrm H}$,

$$S\Lambda^{-1}S^{\mathrm H} = S^*\Lambda S^{\mathrm T}\left(= S\Lambda^* S^{\mathrm H}\right), \qquad \|\delta(\Lambda)\|_\infty = 1\,, \qquad \mathbf{1}^{\mathrm T}|\delta(\Lambda)| = n \tag{1843}$$

characterizes an orthogonal matrix in terms of eigenvalues and eigenvectors.

All permutation matrices Ξ, for example, are nonnegative orthogonal matrices; and *vice versa*. Product or Kronecker product of any permutation matrices remains a permutator. Any product of permutation matrix with orthogonal matrix remains orthogonal. In fact, any product AQ of orthogonal matrices A and Q remains orthogonal by definition. Given any other dimensionally compatible orthogonal matrix U, the mapping $g(A) = U^{\mathrm T}AQ$ is a bijection on the domain of orthogonal matrices (a nonconvex manifold of dimension $\frac{1}{2}n(n-1)$ [56]). [273, §2.1] [274]

[B.9] Orthogonal and unitary matrices are called *unitary linear operators.*

The largest magnitude entry of an orthogonal matrix is 1 ; for each and every $j \in 1 \ldots n$

$$\begin{aligned} \|Q(j\,,:)\|_\infty \leq 1 \\ \|Q(:\,,j)\|_\infty \leq 1 \end{aligned} \tag{1844}$$

Each and every eigenvalue of a (real) orthogonal matrix has magnitude 1 $(\Lambda^{-1} = \Lambda^*)$

$$\lambda(Q) \in \mathbb{C}^n\,, \qquad |\lambda(Q)| = \mathbf{1} \tag{1845}$$

but only the Identity matrix can be simultaneously orthogonal and positive definite. Orthogonal matrices have complex eigenvalues in conjugate pairs: so $\det Q = \pm 1$.

B.5.3 Reflection

A matrix for pointwise reflection is defined by imposing symmetry upon the orthogonal matrix; *id est*, a *reflection matrix* is completely defined by $Q^{-1} = Q^{\mathrm{T}} = Q$. The reflection matrix is a symmetric orthogonal matrix, and *vice versa,* characterized:

$$Q^{\mathrm{T}} = Q\,, \qquad Q^{-1} = Q^{\mathrm{T}}\,, \qquad \|Q\|_2 = 1 \tag{1846}$$

The Householder matrix (§B.3.1) is an example of symmetric orthogonal (reflection) matrix.

Reflection matrices have eigenvalues equal to ± 1 and so $\det Q = \pm 1$. It is natural to expect a relationship between reflection and projection matrices because all projection matrices have eigenvalues belonging to $\{0, 1\}$. In fact, any reflection matrix Q is related to some orthogonal projector P by [235, §1 prob.44]

$$Q = I - 2P \tag{1847}$$

Yet P is, generally, neither orthogonal or invertible. (§E.3.2)

$$\lambda(Q) \in \mathbb{R}^n\,, \qquad |\lambda(Q)| = \mathbf{1} \tag{1848}$$

Whereas P connotes projection on $\mathcal{R}(P)$, here Q connotes reflection with respect to $\mathcal{R}(P)^\perp$.

Matrix $I - 2(I-P) = 2P - I$ represents *antireflection; id est*, reflection about $\mathcal{R}(P)$.

Every orthogonal matrix can be expressed as the product of a rotation and a reflection. The collection of all orthogonal matrices of particular dimension forms a nonconvex set; topologically, it is instead referred to as a *manifold.*

B.5.3.0.1 Example. *Pythagorean sum by antireflection sequence in* $\mathbb{R}^{\mathbf{2}}$.
Figure **186** illustrates a process for determining magnitude of vector $p_0 = [\,x \;\; y\,]^{\mathrm{T}} \in \mathbb{R}^{\mathbf{2}}_+$. The given point p_0 is assumed to be a member of quadrant I with $y \leq x$.[B.10] The idea is to rotate p_0 into alignment with the x axis; its length then becomes equivalent to the rotated x coordinate. Rotation is accomplished by iteration (i index):

First, p_0 is projected on the x axis; $[\,x \;\; 0\,]^{\mathrm{T}}$. Vector $u_0 = [\,x \;\; y/2\,]^{\mathrm{T}}$ is a bisector of difference $p_0 - [\,x \;\; 0\,]^{\mathrm{T}}$, translated to projection $[\,x \;\; 0\,]^{\mathrm{T}}$, and is the perpendicular bisector of chord $\overline{p_0 p_1}$. Point p_1 is the antireflection of p_0 ;

$$p_1 = \left(2\frac{u_0 u_0^{\mathrm{T}}}{u_0^{\mathrm{T}} u_0} - I\right) p_0 \tag{1849}$$

[B.10] p_0 belongs to the monotone nonnegative cone $\mathcal{K}_{\mathcal{M}+} \subset \mathbb{R}^{\mathbf{2}}_+$.

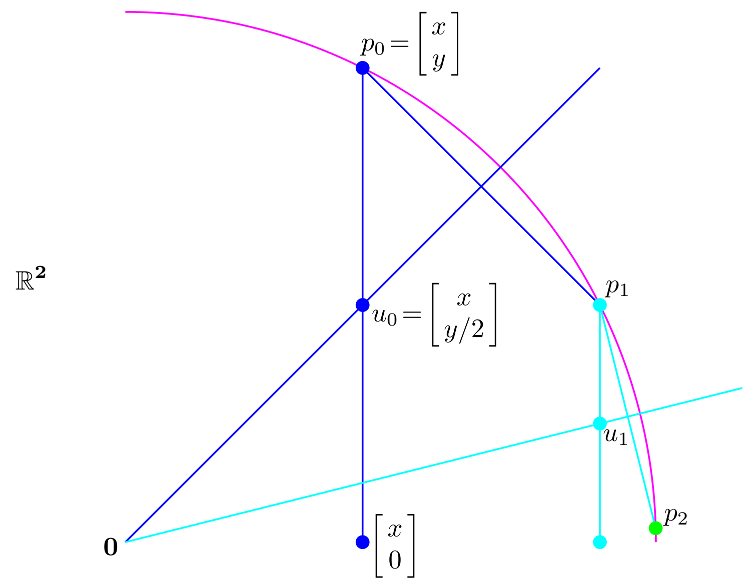


Figure 186: [145] First two of three iterations required, in absence of square root function, to calculate $\sqrt{x^2+y^2}$ to 20 significant digits (presuming infinite precision). Point p_0 is given. u_0 is a bisector of difference $[\,0\;\; y\,]^{\mathrm{T}}$, translated to $[\,x\;\; 0\,]^{\mathrm{T}}$, and is the perpendicular bisector of *chord* $\overline{p_0 p_1}$; similarly, for u_1 . [255, §16/4] Assuming that Moler & Morrison condition [302] $|y| \leq |x|$ holds, iterate p_3 (not shown) is always on quarter circle close to x axis. (p_0 violates Moler & Morrison condition, for illustration. Cartesian axes not drawn.)

In the second iteration, Figure **186** illustrates operation on p_1 and construction of p_2 . Because p_2 is the antireflection of p_1 , each must be equidistant from the perpendicular bisector through u_1 of chord $\overline{p_1 p_2}$.

A third iteration completes the process. Say $p_i = [\,x_i\;\; y_i\,]^{\mathrm{T}}$ with $[\,x_0\;\; y_0\,] \triangleq [\,x\;\; y\,]$ the given point. It is known that the particular selection of bisector $u_i = [\,x_i\;\; y_i/2\,]^{\mathrm{T}}$ diminishes total required number of iterations to three for 20 significant digits ([302] presuming exact arithmetic).

The entire sequence of operations is illustrated programmatically by this Matlab subroutine to calculate magnitude of any vector in $\mathbb{R}^2$ by antireflection:

```
function z = pythag2(x,y)  %z(1)=sqrt(x^2+y^2), z(2)=y coordinate of p3
   if ~[x;y]
      z = [x;y];
   else
      z = sort(abs([x;y]),'descend');
      for i=1:3
         u = [z(1); z(2)/2];
         z = 2*u*(u'*z)/(u'*u) - z;
      end
   end
end
```

Pythagorean sum sans square root was first published in 1983 by Moler & Morrison [302] who did not describe antireflection. □

B.5.3.0.2 Exercise. *Pythagorean sum*: $\|x\in\mathbb{R}^n\|$ *sans square root sans sort.*
MATLAB subroutine `pythag2()` calculates magnitude of any vector in $\mathbb{R}^2$. By constraining perpendicular bisector angle $\measuredangle u_i \leq \frac{\pi}{4}$, show that sorting condition $|y|\leq|x|$ is obviated:

```
function z = pythag2d(x,y)  %z(1)=sqrt(x^2+y^2), z(2)=y coordinate of p4
   if ~[x;y]
      z = [x;y];
   elseif ~x
      z = [abs(y); 0];
   else
      z = abs([x;y]);
      for i=1:4
         u = [z(1); min(z(1), z(2)/2)];
         z = 2*u*(u'*z)/(u'*u) - z;
      end
   end
end
```

Elimination of `sort()` incurred an extra iteration in $\mathbb{R}^2$.

Antireflection admits simple modifications that enable calculation of Euclidean norm for any vector in $\mathbb{R}^n$. Show that Euclidean norm may be calculated with only a small number of iterations in absence of square root and sorting functions:

```
function z = pythag3d(x)  %z(1)=||x||, z(2:end)=coordinates of p13
   if ~x
      z = x;
   elseif ~x(1)
      z = [pythag3d(x(2:end)); 0];
   else
      z = abs(x);
      for i=1:13
         u = [z(1); min(z(1), z(2:end)/2)];
         z = 2*u*(u'*z)/(u'*u) - z;
      end
   end
end
```

Projection is on the first nonzero coordinate axis, as in §B.5.3.0.1, then antireflection is about a vector u_i normal to a hyperplane. Demonstrate that number of required iterations does not grow linearly with n; contrary to [302, §3], growth is much slower. ▼

B.5.4 Rotation of range and rowspace

Given orthogonal matrix Q, column vectors of a matrix X are simultaneously rotated about the origin via product QX. In three dimensions ($X\in\mathbb{R}^{3\times N}$), precise meaning of rotation is best illustrated in Figure **187** where a gimbal aids visualization of what is achievable; mathematically, (§5.5.2.0.1)

$$Q = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1850}$$

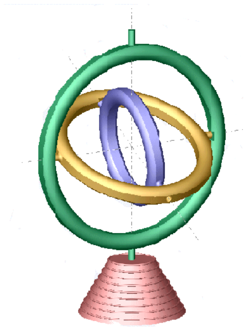


Figure 187: *Gimbal*: a mechanism imparting three degrees of dimensional freedom to a Euclidean body suspended at its center. Each ring is free to rotate about one axis. (Drawing by courtesy of The MathWorks Inc.)

B.5.4.0.1 Example. *One axis of revolution.*
Partition $n+1$-dimensional Euclidean space $\mathbb{R}^{n+1} \triangleq \begin{bmatrix} \mathbb{R}^n \\ \mathbb{R} \end{bmatrix}$ and define an n-dimensional subspace

$$\mathcal{R} \triangleq \{\lambda \in \mathbb{R}^{n+1} \mid \mathbf{1}^{\mathrm{T}}\lambda = 0\} \tag{1851}$$

(a hyperplane through the origin). We want an orthogonal matrix that rotates a list in the columns of matrix $X \in \mathbb{R}^{n+1\times N}$ through the dihedral angle between $\mathbb{R}^n$ and $\mathcal{R}$ (§2.4.3)

$$\measuredangle(\mathbb{R}^n,\, \mathcal{R}) \;=\; \arccos\!\left(\frac{\langle e_{n+1}\,,\, \mathbf{1}\rangle}{\|e_{n+1}\|\;\|\mathbf{1}\|}\right) \;=\; \arccos\!\left(\frac{1}{\sqrt{n+1}}\right) \text{ radians} \tag{1852}$$

The vertex-description of the nonnegative orthant in $\mathbb{R}^{n+1}$ is

$$\{[\,e_1 \;\; e_2 \;\cdots\; e_{n+1}\,]\,a \mid a \succeq 0\} \;=\; \{a \succeq 0\} \;=\; \mathbb{R}^{n+1}_+ \subset \mathbb{R}^{n+1} \tag{1853}$$

Consider rotation of these vertices via orthogonal matrix

$$Q \triangleq [\,\mathbf{1}\tfrac{1}{\sqrt{n+1}} \quad \Xi V_{\mathcal{W}}\,]\,\Xi \in \mathbb{R}^{n+1\times n+1} \tag{1854}$$

where permutation matrix $\Xi \in \mathbb{S}^{n+1}$ is defined in (1920) and where $V_{\mathcal{W}} \in \mathbb{R}^{n+1\times n}$ is the orthonormal auxiliary matrix defined in §B.4.3. This particular orthogonal matrix is selected because it rotates any point in subspace $\mathbb{R}^n$ about one axis of revolution onto $\mathcal{R}$; *e.g*, rotation Qe_{n+1} aligns the last standard basis vector with subspace normal $\mathcal{R}^{\perp} = \mathbf{1}$. The rotated standard basis vectors remaining are orthonormal spanning $\mathcal{R}$. □

Another interpretation of product QX is rotation/reflection of $\mathcal{R}(X)$. Rotation of X as in QXQ^{T} is a simultaneous rotation/reflection of range and rowspace.[B.11]

Proof. Any matrix can be expressed as a singular value decomposition $X = U\Sigma W^{\mathrm{T}}$ (1734) where $\delta^2(\Sigma) = \Sigma$, $\mathcal{R}(U) \supseteq \mathcal{R}(X)$, and $\mathcal{R}(W) \supseteq \mathcal{R}(X^{\mathrm{T}})$. ◆

[B.11] Product $Q^{\mathrm{T}}AQ$ can be regarded as coordinate transformation; *e.g*, given linear map $y = Ax : \mathbb{R}^n \to \mathbb{R}^n$ and orthogonal Q, the transformation $Qy = AQx$ is a rotation/reflection of range and rowspace (145) of matrix A where $Qy \in \mathcal{R}(A)$ and $Qx \in \mathcal{R}(A^{\mathrm{T}})$ (146).

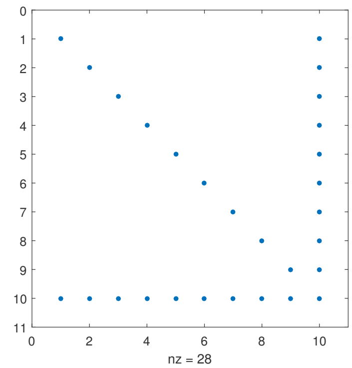


Figure 188: 10×10 arrow matrix. Twenty eight nonzero (nz) entries indicated.

B.5.5 Matrix rotation

Orthogonal matrices are also employed to rotate/reflect other matrices like vectors: [185, §12.4.1] Given orthogonal matrix Q, the product $Q^{\mathrm T}A$ will rotate $A\in\mathbb{R}^{n\times n}$ in the Euclidean sense in $\mathbb{R}^{n^2}$ because Frobenius' norm is orthogonally invariant (§2.2.1);

$$\|Q^{\mathrm T}A\|_{\mathrm F} \;=\; \sqrt{\operatorname{tr}(A^{\mathrm T}QQ^{\mathrm T}A)} \;=\; \|A\|_{\mathrm F} \tag{1855}$$

(likewise for AQ). Were A symmetric, such a rotation would depart from $\mathbb{S}^n$. One remedy is to instead form product $Q^{\mathrm T}A\,Q$ because

$$\|Q^{\mathrm T}A\,Q\|_{\mathrm F} \;=\; \sqrt{\operatorname{tr}(Q^{\mathrm T}A^{\mathrm T}QQ^{\mathrm T}A\,Q)} \;=\; \|A\|_{\mathrm F} \tag{1856}$$

By §A.1.1 *no.*33,

$$\operatorname{vec} Q^{\mathrm T}A\,Q = (Q\otimes Q)^{\mathrm T}\operatorname{vec} A \tag{1857}$$

which is a rotation of the vectorized A matrix because Kronecker product of any orthogonal matrices remains orthogonal; *e.g*, by §A.1.1 *no.*43,

$$(Q\otimes Q)^{\mathrm T}(Q\otimes Q) = I \tag{1858}$$

Matrix A is *orthogonally equivalent* to B if $B = S^{\mathrm T}A\,S$ for some orthogonal matrix S. Every square matrix, for example, is orthogonally equivalent to a matrix having equal entries along the main diagonal. [233, §2.2, prob.3]

B.6 Arrow matrix

Consider a partitioned symmetric $n\times n$-dimensional matrix A that has *arrow* [330] (or *arrowhead* [373]) form constituted by vectors $a\,,b\in\mathbb{R}^{n-1}$ and real scalar c:

$$A \;\triangleq\; \begin{bmatrix} \delta(a) & b \\ b^{\mathrm T} & c \end{bmatrix} \in \mathbb{S}^n \tag{1859}$$

Figure **188** illustrates sparsity pattern of an arrow matrix. Embedding of diagonal matrix $\delta(a)$ makes relative sparsity increasing with dimension. Because an arrow matrix is a kind of bordered matrix, eigenvalues of $\delta(a)$ and A are interlaced;

$$\lambda_n \le (\Xi^{\mathrm T}a)_{n-1} \le\; \lambda_{n-1} \le (\Xi^{\mathrm T}a)_{n-2} \le \cdots \le (\Xi^{\mathrm T}a)_1 \le\; \lambda_1 \tag{1860}$$

[374, §6.4] [233, §4.3] [370, §IV.4.1] denoting nonincreasingly ordered eigenvalues of A by vector $\lambda\in\mathbb{R}^n$, and those of $\delta(a)$ by $\Xi^{\mathrm T}a\in\mathbb{R}^{n-1}$ where Ξ is a permutation matrix arranging a into nonincreasing order: $\delta(a) = \Xi\delta(\Xi^{\mathrm T}a)\Xi^{\mathrm T}$ (§A.5.1.2).

B.6.1 positive semidefinite arrow matrix

i) Nonnegative main diagonal $a \succeq 0$ insures $n-1$ nonnegative eigenvalues in (1860).

Positive semidefiniteness is left determined by smallest eigenvalue λ_n :

$$A \succeq 0 \quad \begin{array}{lll} \Leftrightarrow \;\; a \succeq 0\,, & b^{\mathrm T}\big(I - \delta(a)\delta(a)^\dagger\big) = \mathbf{0}\,, & c - b^{\mathrm T}\delta(a)^\dagger b \geq 0 \\ \Leftrightarrow \;\; c \geq 0\,, & b(1 - cc^\dagger) = \mathbf{0}\,, & \delta(a) - c^\dagger b b^{\mathrm T} \succeq 0 \end{array} \tag{1861}$$

Schur complement condition (§A.4) $b^{\mathrm T}\big(I - \delta(a)\delta(a)^\dagger\big) = \mathbf{0}$ is most simply a requirement for

ii) a zero entry in vector b wherever there is a corresponding zero entry in vector a .

In other words, vector b can reside anywhere in a Cartesian subspace of $\mathbb{R}^{n-1}$ that is determined solely by indices of the nonzero entries in vector a .

iii) $c \geq b^{\mathrm T}\delta(a)^\dagger b$ provides a tight lower bound for scalar c .

As shown in §3.5.1, $b^{\mathrm T}\delta(a)^\dagger b$ is simultaneously convex in vectors a and b .