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.*

−Marie A. Vitulli [251]

Citation: Jon Dattorro, *Convex Optimization & Euclidean Distance Geometry,* Meboo Publishing USA, 2005.

B.1 Rank-one matrix (dyad)

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

$$\Psi = uv^T \in \mathbb{R}^{M\times N} \tag{1236}$$

where $u\in\mathbb{R}^M$ and $v\in\mathbb{R}^N$, is rank-one and called a *dyad.* Conversely, any rank-one matrix must have the form Ψ. [133, prob.1.4.1] The product $-uv^T$ is a *negative dyad.* For matrix products AB^T, in general, we have

$$\mathcal{R}(AB^T) \subseteq \mathcal{R}(A)\ , \qquad \mathcal{N}(AB^T) \supseteq \mathcal{N}(B^T) \tag{1237}$$

with equality when $B = A$ [223, §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

$$\mathcal{R}(uv^T) = \mathcal{R}(u)\ , \qquad \mathcal{N}(uv^T) = \mathcal{N}(v^T) \equiv v^\perp \tag{1238}$$

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

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

$$\Psi = uv^T = \mathbf{0} \;\Leftrightarrow\; u=\mathbf{0} \;\text{ or }\; v=\mathbf{0} \tag{1239}$$

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

$$\|\Psi\|_2 \;=\; \|uv^T\|_F \;=\; \|uv^T\|_2 \;=\; \|u\|\,\|v\| \tag{1240}$$

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

$$\Psi^\dagger \;=\; (uv^T)^\dagger \;=\; \frac{vu^T}{\|u\|^2\,\|v\|^2} \tag{1241}$$

When dyad $uv^T\in\mathbb{R}^{N\times N}$ 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

$$\lambda \;=\; v^Tu \;=\; \operatorname{tr}(uv^T) \in \mathbb{R} \tag{1242}$$

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

$$\begin{aligned}\mathcal{R}(AA^T) \;&=\; \{AA^Ty \mid y\in\mathbb{R}^m\}\\ &\supseteq\; \{AA^Ty \mid A^Ty\in\mathcal{R}(A^T)\} = \mathcal{R}(A)\ \text{ by (115)}\end{aligned}$$ ♦

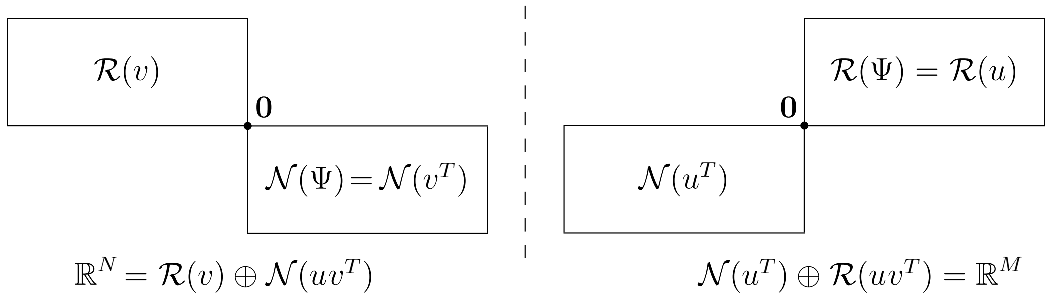


Figure 104: The four fundamental subspaces [225, §3.6] of any dyad $\Psi = uv^T \in \mathbb{R}^{M\times N}$. $\Psi(x) \triangleq uv^T x$ is a linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$. The map from $\mathcal{R}(v)$ to $\mathcal{R}(u)$ is bijective. [223, §3.1]

The determinant is the product of the eigenvalues; so, it is always true that

$$\det \Psi \;=\; \det(uv^T) \;=\; 0 \tag{1243}$$

When $\lambda = 1$, the square dyad is a nonorthogonal projector projecting on its range ($\Psi^2 = \Psi$, §E.1). 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

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} \;=\; \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \tag{1244}$$

for example, has two 0-eigenvalues. In other words, the value of 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, $\dim u = M = \dim v = N$:

Proof. Figure **104** shows the four fundamental subspaces for the dyad. Linear operator $\Psi : \mathbb{R}^N \to \mathbb{R}^M$ provides a map between vector spaces that remain distinct when $M = N$;

$$\begin{aligned} u &\in \mathcal{R}(uv^T) \\ u \in \mathcal{N}(uv^T) &\Leftrightarrow v^T u = 0 \\ \mathcal{R}(uv^T) \cap \mathcal{N}(uv^T) &= \emptyset \end{aligned} \tag{1245}$$

♦

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

B.1.0.1 rank-one modification

If $A \in \mathbb{R}^{N\times N}$ is any nonsingular matrix and $1+v^TA^{-1}u \neq 0$, then [145, App.6] [272, §2.3, prob.16] [89, §4.11.2] (Sherman-Morrison)

$$(A+uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u} \tag{1246}$$

B.1.0.2 dyad symmetry

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

$$uv^T \succeq 0 \;\Leftrightarrow\; v=u \tag{1247}$$

and the remaining eigenvalue is almost always positive;

$$\lambda \;=\; u^Tu \;=\; \mathrm{tr}(uu^T) > 0 \;\text{ unless }\; u=\mathbf{0} \tag{1248}$$

The matrix

$$\begin{bmatrix} \Psi & u \\ u^T & 1 \end{bmatrix} \tag{1249}$$

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

B.1.1 Dyad independence

Now we consider a sum of dyads like (1236) as encountered in diagonalization and singular value decomposition:

$$\mathcal{R}\left(\sum_{i=1}^k s_i w_i^T\right) = \sum_{i=1}^k \mathcal{R}\big(s_i w_i^T\big) = \sum_{i=1}^k \mathcal{R}(s_i) \;\Leftarrow\; w_i \;\forall\, i \text{ are l.i.} \tag{1250}$$

range of the summation is the vector sum of ranges.[B.3] (Theorem B.1.1.1.1) Under the assumption the dyads are linearly independent (l.i.), then the vector sums are unique (p.645): for $\{w_i\}$ l.i. and $\{s_i\}$ l.i.

$$\mathcal{R}\left(\sum_{i=1}^k s_i w_i^T\right) = \mathcal{R}\big(s_1 w_1^T\big) \oplus \ldots \oplus \mathcal{R}\big(s_k w_k^T\big) = \mathcal{R}(s_1) \oplus \ldots \oplus \mathcal{R}(s_k) \tag{1251}$$

[B.3] Move of range $\mathcal{R}$ to inside the summation depends on linear independence of $\{w_i\}$.

B.1.1.0.1 Definition. *Linearly independent dyads.* [139, p.29, thm.11] [230, p.2] The set of k dyads

$$\left\{s_i w_i^T \mid i=1\ldots k\right\} \tag{1252}$$

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

$$\operatorname{rank}\left(SW^T \triangleq \sum_{i=1}^{k} s_i w_i^T\right) = k \tag{1253}$$

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

As defined, dyad independence does not preclude existence of a nullspace $\mathcal{N}(SW^T)$, nor does it imply SW^T is full-rank. In absence of an assumption of independence, generally, $\operatorname{rank} SW^T\leq k$. Conversely, any rank-$k$ matrix can be written in the form SW^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 $\{s_i w_i^T\in\mathbb{C}^{M\times N},\ i=1\ldots k\}$ are l.i. ⋄

Proof. Linear independence of k dyads is identical to definition (1253).
(⇒) Suppose $\{s_i\}$ and $\{w_i\}$ are each linearly independent sets. Invoking Sylvester's rank inequality, [133, §0.4] [272, §2.4]

$$\operatorname{rank} S+\operatorname{rank} W-k \;\leq\; \operatorname{rank}(SW^T) \;\leq\; \min\{\operatorname{rank} S\,,\ \operatorname{rank} W\} \;\;(\leq\; k) \tag{1254}$$

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

$$k\leq\min\{\operatorname{rank} S\,,\ \operatorname{rank} W\}\leq k \tag{1255}$$

implying the vector sets are each independent. ♦

B.1.1.1 Biorthogonality condition, Range and Nullspace of Sum

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

$$W^TS = I \quad\Leftrightarrow\quad \operatorname{rank} S = \operatorname{rank} W = k \leq M = N \tag{1256}$$

To see that, we need only show: $\mathcal{N}(S) = \mathbf{0} \;\Leftrightarrow\; \exists\, B \;\ni\; BS = I$.[B.4]
(⇐) Assume $BS = I$. Then $\mathcal{N}(BS) = \mathbf{0} = \{x \mid BSx = \mathbf{0}\} \supseteq \mathcal{N}(S)$. (1237)
(⇒) If $\mathcal{N}(S) = \mathbf{0}$ then S must be full-rank skinny-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^T here. Because of reciprocity with S, it immediately follows: $\mathcal{N}(W) = \mathbf{0} \;\Leftrightarrow\; \exists\, S \;\ni\; S^TW = I$. ♦

Dyads produced by diagonalization, for example, are independent because of their inherent biorthogonality. (§A.5.1) 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^T where $S \in \mathbb{C}^{M\times k}$ and $W \in \mathbb{C}^{N\times k}$

$$\begin{aligned} \mathcal{N}(SW^T) = \mathcal{N}(W^T) \quad&\Leftarrow\quad \exists\, B \;\ni\; BS = I \\ \mathcal{R}(SW^T) = \mathcal{R}(S) \quad&\Leftarrow\quad \exists\, Z \;\ni\; W^TZ = I \end{aligned} \tag{1257}$$

⋄

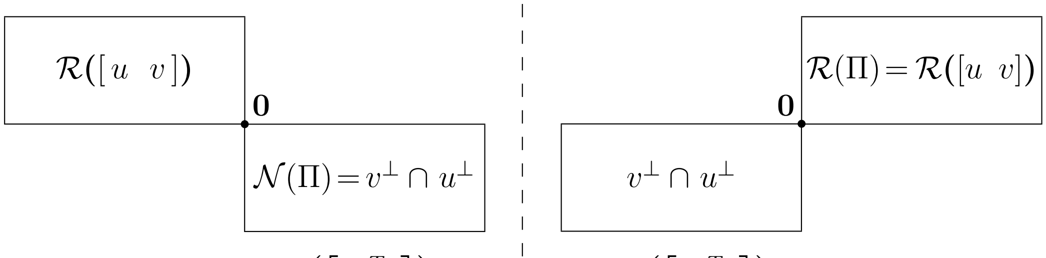
Proof. **(⇒)** $\mathcal{N}(SW^T) \supseteq \mathcal{N}(W^T)$ and $\mathcal{R}(SW^T) \subseteq \mathcal{R}(S)$ are obvious.
(⇐) Assume the existence of a left inverse $B \in \mathbb{R}^{k\times N}$ and a right inverse $Z \in \mathbb{R}^{N\times k}$.[B.5]

$$\mathcal{N}(SW^T) \;=\; \{x \mid SW^Tx = \mathbf{0}\} \;\subseteq\; \{x \mid BSW^Tx = \mathbf{0}\} \;=\; \mathcal{N}(W^T) \tag{1258}$$

$$\mathcal{R}(SW^T) \;=\; \{SW^Tx \mid x \in \mathbb{R}^N\} \;\supseteq\; \{SW^TZy \mid Zy \in \mathbb{R}^N\} \;=\; \mathcal{R}(S) \tag{1259}$$

♦

[B.4] Left inverse is not unique, in general.
[B.5] By counter example, the theorem's converse cannot be true; *e.g.*, $S = W = [\,\mathbf{1} \;\; \mathbf{0}\,]$.



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

Figure 105: Four fundamental subspaces [225, §3.6] of a doublet $\Pi = uv^T + vu^T \in \mathbb{S}^N$. $\Pi(x) = (uv^T + vu^T)x$ is a linear bijective mapping from $\mathcal{R}([\,u \;\; v\,])$ to $\mathcal{R}([\,u \;\; v\,])$.

B.2 Doublet

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

$$\Pi \;=\; uv^T + vu^T \;=\; [\,u \;\; v\,] \begin{bmatrix} v^T \\ u^T \end{bmatrix} \;=\; SW^T \in \mathbb{S}^N \tag{1260}$$

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

$$\Pi = uv^T + vu^T = \mathbf{0} \;\Leftrightarrow\; u = \mathbf{0} \;\text{ or }\; v = \mathbf{0} \tag{1261}$$

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

$$\lambda_1 \triangleq u^Tv + \|uv^T\|\;, \qquad \lambda_2 \triangleq u^Tv - \|uv^T\| \tag{1262}$$

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

$$x_1 \triangleq \frac{u}{\|u\|} + \frac{v}{\|v\|}\;, \qquad x_2 \triangleq \frac{u}{\|u\|} - \frac{v}{\|v\|} \tag{1263}$$

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.

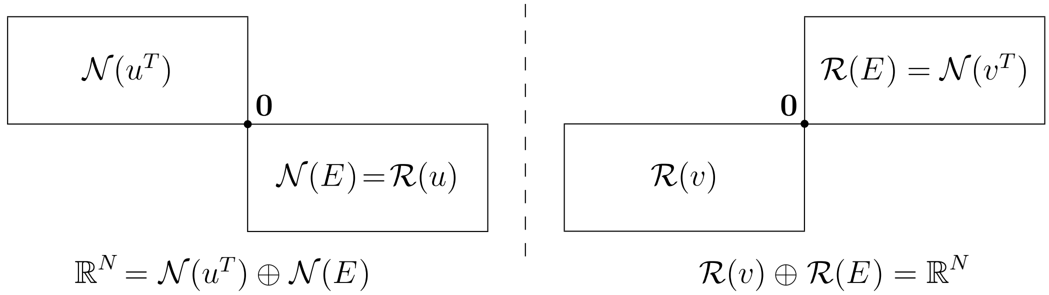


Figure 106: $v^Tu = 1/\zeta$. The four fundamental subspaces [225, §3.6] of elementary matrix E as a linear mapping $E(x) = \left(I - \frac{uv^T}{v^Tu}\right)x$.

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

$$\mathcal{R}(\Pi) = \mathcal{R}([\,u \;\; v\,]) \;, \qquad \mathcal{N}(\Pi) = v^\perp \cap u^\perp \tag{1264}$$

of respective dimension 2 and $N-2$.

B.3 Elementary matrix

A matrix of the form

$$E = I - \zeta uv^T \in \mathbb{R}^{N\times N} \tag{1265}$$

where $\zeta \in \mathbb{R}$ is finite and $u,v \in \mathbb{R}^N$, is called an *elementary matrix* or a *rank-one modification of the identity*. [135] 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

$$\lambda = 1 - \zeta v^Tu \tag{1266}$$

corresponds to eigenvector u .[B.6] From [145, App.7.A.26] the determinant:

$$\det E \;=\; 1 - \operatorname{tr}\!\left(\zeta uv^T\right) \;=\; \lambda \tag{1267}$$

[B.6] 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.

If $\lambda \neq 0$ then E is invertible; [89]

$$E^{-1} = I + \frac{\zeta}{\lambda} uv^T \tag{1268}$$

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. (§A.7.3.0.1) The nullspace exists, therefore, when $\lambda = 0$; *id est*, when $v^T u = 1/\zeta$, rather, whenever u belongs to the hyperplane $\{z \in \mathbb{R}^N \mid v^T z = 1/\zeta\}$. 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 (1265) 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,

$$\mathcal{R}(E) = \mathcal{N}(v^T), \qquad \mathcal{N}(E) = \mathcal{R}(u), \qquad v^T u = 1/\zeta \tag{1269}$$

illustrated in Figure **106**, which is opposite to the assignment of subspaces for a dyad (Figure **104**). Otherwise, $\mathcal{R}(E) = \mathbb{R}^N$.

When $E = E^T$, the spectral norm is

$$\|E\|_2 = \max\{1\,,\,|\lambda|\} \tag{1270}$$

B.3.1 Householder matrix

An elementary matrix is called a Householder matrix when it has the defining form, for nonzero vector u [95, §5.1.2] [89, §4.10.1] [223, §7.3] [133, §2.2]

$$H = I - 2\frac{uu^T}{u^T u} \in \mathbb{S}^N \tag{1271}$$

which is a symmetric orthogonal (reflection) matrix $\big(H^{-1} = H^T = H$ (§B.5.2)$\big)$. 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

$$\det H = -1 \tag{1272}$$

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,

$$H^T = H\,, \qquad H^{-1} = H^T\,, \qquad \|H\|_2 = 1\,, \qquad H \not\succeq 0 \tag{1273}$$

For example, the permutation matrix

$$\Xi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{1274}$$

is a Householder matrix having $u = [\,0\;\;1\;\;-1\,]^T/\sqrt{2}$. Not all permutation matrices are Householder matrices, although all permutation matrices are orthogonal matrices. [223, §3.4] Neither are all symmetric permutation matrices Householder matrices; *e.g.*, $\Xi = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ (1356) 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 the geometric center α_c (§4.5.1.0.1) of some list X to the origin. In place of $X - \alpha_c \mathbf{1}^T$ we may write XV as in (548) where

$$V \;\triangleq\; I - \frac{1}{N}\mathbf{1}\mathbf{1}^T \in \mathbb{S}^N \tag{491}$$

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 $V = V^T$ is diagonalizable. Because

$$V\mathbf{1} = \mathbf{0} \tag{1275}$$

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

Because

$$V^2 = V \tag{1276}$$

and $V^T = V$, elementary matrix V is also a projection matrix (§E.3) projecting orthogonally on its range $\mathcal{N}(\mathbf{1}^T)$.

$$V = I - \mathbf{1}(\mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T \tag{1277}$$

The $\{0,1\}$ eigenvalues also indicate diagonalizable V is a projection matrix. [272, §4.1, thm.4.1] Symmetry of V denotes orthogonal projection; from (1519),

$$V^T = V\ , \qquad V^\dagger = V\ , \qquad \|V\|_2 = 1\,, \qquad V \succeq 0 \tag{1278}$$

Matrix V is also circulant [103].

B.4.1.0.1 Example. *Relationship of auxiliary to Householder matrix.* Let $H \in \mathbb{S}^N$ be a Householder matrix (1271) defined by

$$u = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1+\sqrt{N} \end{bmatrix} \in \mathbb{R}^N \tag{1279}$$

Then we have [92, §2]

$$V = H \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} H \tag{1280}$$

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

$$-HDH \triangleq -\begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \tag{1281}$$

where b is a vector. Then because H is nonsingular (§A.3.1.0.5) [116, §3]

$$-VDV = -H \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} H \succeq 0 \;\Leftrightarrow\; -A \succeq 0 \tag{1282}$$

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. $V_{\mathcal{N}} = \frac{1}{\sqrt{2}}\begin{bmatrix} -\mathbf{1}^T \\ I \end{bmatrix} \in \mathbb{R}^{N\times N-1}$
2. $V_{\mathcal{N}}^T\mathbf{1} = \mathbf{0}$
3. $I - e_1\mathbf{1}^T = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$
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. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^\dagger = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & I \end{bmatrix} V$
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. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}^\dagger\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & I \end{bmatrix}$
12. $\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}\begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & I \end{bmatrix} = \begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix}$
13. $\begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & I \end{bmatrix}\begin{bmatrix}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{bmatrix} = \begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & I \end{bmatrix}$
14. $[\,V_{\mathcal{N}} \quad \frac{1}{\sqrt{2}}\mathbf{1}\,]^{-1} = \begin{bmatrix} V_{\mathcal{N}}^\dagger \\ \frac{\sqrt{2}}{N}\mathbf{1}^T \end{bmatrix}$
15. $V_{\mathcal{N}}^\dagger = \sqrt{2}\begin{bmatrix}-\frac{1}{N}\mathbf{1} & I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\end{bmatrix} \in \mathbb{R}^{N-1\times N}, \qquad \left(I - \frac{1}{N}\mathbf{1}\mathbf{1}^T \in \mathbb{S}^{N-1}\right)$
16. $V_{\mathcal{N}}^\dagger\mathbf{1} = \mathbf{0}$
17. $V_{\mathcal{N}}^\dagger V_{\mathcal{N}} = I$

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

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

20. $D = [d_{ij}] \in \mathbb{S}_h^N$ (493)

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

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

$$E = k_1 I - k_2 \mathbf{1}\mathbf{1}^T \tag{1283}$$

where $k_1, k_2 \in \mathbb{R}$,[B.7] will make $\operatorname{tr}(-ED)$ proportional to $\sum d_{ij}$.

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

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

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

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

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 $V_{\mathcal{W}}$

The skinny matrix

$$V_{\mathcal{W}} \triangleq \begin{bmatrix} \frac{-1}{\sqrt{N}} & \frac{-1}{\sqrt{N}} & \cdots & \frac{-1}{\sqrt{N}} \\ 1 + \frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & \frac{-1}{N+\sqrt{N}} \\ \frac{-1}{N+\sqrt{N}} & \ddots & \ddots & \frac{-1}{N+\sqrt{N}} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & 1 + \frac{-1}{N+\sqrt{N}} \end{bmatrix} \in \mathbb{R}^{N\times N-1} \tag{1284}$$

[B.7] 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. [200]

has $\mathcal{R}(V_{\mathcal{W}})=\mathcal{N}(\mathbf{1}^T)$ and orthonormal columns. [4] We defined three auxiliary V-matrices: V, $V_{\mathcal{N}}$ (474), and $V_{\mathcal{W}}$ sharing some attributes listed in Table **B.4.4**. For example, V can be expressed

$$V = V_{\mathcal{W}} V_{\mathcal{W}}^T = V_{\mathcal{N}} V_{\mathcal{N}}^\dagger \tag{1285}$$

but $V_{\mathcal{W}}^T V_{\mathcal{W}} = I$ means V is an orthogonal projector (1516) and

$$V_{\mathcal{W}}^\dagger = V_{\mathcal{W}}^T\,, \qquad \|V_{\mathcal{W}}\|_2 = 1\,, \qquad V_{\mathcal{W}}^T\mathbf{1} = \mathbf{0} \tag{1286}$$

B.4.4 Auxiliary V-matrix Table

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

B.4.5 More auxiliary matrices

Mathar shows [171, §2] that any elementary matrix (§B.3) of the form

$$V_{\mathcal{M}} = I - b\mathbf{1}^T \in \mathbb{R}^{N\times N} \tag{1287}$$

such that $b^T\mathbf{1}=1$ (*confer* [97, §2]), is an auxiliary V-matrix having

$$\begin{array}{ll} \mathcal{R}(V_{\mathcal{M}}^T) = \mathcal{N}(b^T)\,, & \mathcal{R}(V_{\mathcal{M}}) = \mathcal{N}(\mathbf{1}^T) \\ \mathcal{N}(V_{\mathcal{M}}) = \mathcal{R}(b)\,, & \mathcal{N}(V_{\mathcal{M}}^T) = \mathcal{R}(\mathbf{1}) \end{array} \tag{1288}$$

Given $X\in\mathbb{R}^{n\times N}$, the choice $b=\frac{1}{N}\mathbf{1}$ ($V_{\mathcal{M}}=V$) minimizes $\|X(I-b\mathbf{1}^T)\|_{\mathrm{F}}$. [99, §3.2.1]

B.5 Orthogonal matrix

B.5.1 Vector rotation

The property $Q^{-1} = Q^T$ completely defines an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ employed to effect vector rotation; [223, §2.6, §3.4] [225, §6.5] [133, §2.1] for $x \in \mathbb{R}^n$

$$\|Qx\| = \|x\| \tag{1289}$$

The orthogonal matrix is characterized:

$$Q^{-1} = Q^T\,, \qquad \|Q\|_2 = 1 \tag{1290}$$

Applying characterization (1290) to Q^T we see it too is an orthogonal matrix. Hence the rows and columns of Q respectively form an orthonormal set.

All permutation matrices Ξ , for example, are orthogonal matrices. The largest magnitude entry of any orthogonal matrix is 1; for each and every $j \in 1 \ldots n$

$$\begin{aligned} \|Q(j\,,:)\|_\infty &\leq 1 \\ \|Q(:\,,j)\|_\infty &\leq 1 \end{aligned} \tag{1291}$$

Each and every eigenvalue of a (real) orthogonal matrix has magnitude 1

$$\lambda(Q) \in \mathbb{C}^n\,, \qquad |\lambda(Q)| = \mathbf{1} \tag{1292}$$

while only the identity matrix can be simultaneously positive definite and orthogonal.

A *unitary matrix* is a complex generalization of the orthogonal matrix. The conjugate transpose defines it: $U^{-1} = U^H$. An orthogonal matrix is simply a real unitary matrix.

B.5.2 Reflection

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

$$Q^T = Q\,, \qquad Q^{-1} = Q^T\,, \qquad \|Q\|_2 = 1 \tag{1293}$$

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

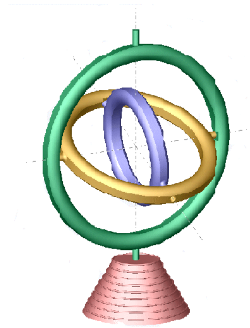


Figure 107: *Gimbal*: a mechanism imparting three degrees of dimensional freedom to a Euclidean body suspended at the device's center. Each ring is free to rotate about one axis. (Drawing courtesy of The MathWorks Inc.)

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 [135, §1, prob.44]

$$Q = I - 2P \tag{1294}$$

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

$$\lambda(Q) \in \mathbb{R}^n\,, \qquad |\lambda(Q)| = \mathbf{1} \tag{1295}$$

Reflection is with respect to $\mathcal{R}(P)^\perp$. Matrix $2P-I$ represents antireflection.

Every orthogonal matrix can be expressed as the product of a rotation and a reflection. The collection of all orthogonal matrices of particular dimension does not form a convex set.

B.5.3 Rotation of range and rowspace

Given orthogonal matrix Q, column vectors of a matrix X are simultaneously rotated by the product QX. In three dimensions ($X \in \mathbb{R}^{\mathbf{3}\times N}$), the precise meaning of rotation is best illustrated in Figure **107** where the gimbal aids visualization of rotation achievable about the origin.

B.5.3.0.1 Example. *One axis of revolution.*
Partition an $n+1$-dimensional Euclidean space $\mathbb{R}^{n+1} \triangleq \begin{bmatrix} \mathbb{R}^n \\ \mathbb{R} \end{bmatrix}$ and define an n-dimensional subspace

$$\mathcal{R} \triangleq \{\lambda \in \mathbb{R}^{n+1} \mid \mathbf{1}^T\lambda = 0\} \tag{1296}$$

(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 $\mathcal{R}$: $\measuredangle(\mathbb{R}^n, \mathcal{R}) = \arccos\big(1/\sqrt{n+1}\big)$ radians. The vertex-description of the nonnegative orthant in $\mathbb{R}^{n+1}$ is

$$\{[\, e_1 \;\; e_2 \cdots e_{n+1}\,]\, a \mid a \succeq 0\} = \{a \succeq 0\} \subset \mathbb{R}^{n+1} \tag{1297}$$

Consider rotation of these vertices via orthogonal matrix

$$Q \triangleq [\, \mathbf{1}\tfrac{1}{\sqrt{n+1}} \quad \Xi V_{\mathcal{W}}\,]\,\Xi \in \mathbb{R}^{n+1\times n+1} \tag{1298}$$

where permutation matrix $\Xi \in \mathbb{S}^{n+1}$ is defined in (1356), and $V_{\mathcal{W}} \in \mathbb{R}^{n+1\times n}$ is the orthonormal auxiliary matrix defined in §B.4.3. This particular orthogonal matrix is selected because it rotates any point in $\mathbb{R}^n$ about one axis of revolution onto $\mathcal{R}$; *e.g.*, rotation Qe_{n+1} aligns the last standard basis vector with subspace normal $\mathcal{R}^\perp = \mathbf{1}$, and from these two vectors we get $\measuredangle(\mathbb{R}^n, \mathcal{R})$. 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 the simultaneous rotation/reflection of range and rowspace.[B.8]

Proof. Any matrix can be expressed as a singular value decomposition $X = U\Sigma W^T$ (1195) where $\delta^2(\Sigma) = \Sigma$, $\mathcal{R}(U) \supseteq \mathcal{R}(X)$, and $\mathcal{R}(W) \supseteq \mathcal{R}(X^T)$. ♦

[B.8] The product Q^TAQ can be regarded as a 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 the range and rowspace (114) of matrix A where $Qy \in \mathcal{R}(A)$ and $Qx \in \mathcal{R}(A^T)$ (115).

B.5.4 Matrix rotation

Orthogonal matrices are also employed to rotate/reflect like vectors other matrices: [*sic*] [95, §12.4.1] Given orthogonal matrix Q, the product Q^TA will rotate $A\in\mathbb{R}^{n\times n}$ in the Euclidean sense in $\mathbb{R}^{n^2}$ because the Frobenius norm is orthogonally invariant (§2.2.1);

$$\|Q^TA\|_F = \sqrt{\operatorname{tr}(A^TQQ^TA)} = \|A\|_F \tag{1299}$$

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

$$\|Q^TAQ\|_F = \sqrt{\operatorname{tr}(Q^TA^TQQ^TAQ)} = \|A\|_F \tag{1300}$$

Matrix A is *orthogonally equivalent* to B if $B=S^TAS$ for some orthogonal matrix S. Every square matrix, for example, is orthogonally equivalent to a matrix having equal entries along the main diagonal. [133, §2.2, prob.3]

B.5.4.1 bijection

Any product of orthogonal matrices AQ remains orthogonal. Given any other dimensionally compatible orthogonal matrix U, the mapping $g(A)=U^TAQ$ is a linear bijection on the domain of orthogonal matrices. [158, §2.1]