Appendix A

Linear algebra

A.1 Main-diagonal δ operator, λ , tr , vec

We introduce notation δ denoting the main-diagonal linear selfadjoint operator. When linear function δ operates on a square matrix $A\in\mathbb{R}^{N\times N}$, $\delta(A)$ returns a vector composed of all the entries from the main diagonal in the natural order;

$$\delta(A) \in \mathbb{R}^N \tag{1504}$$

Operating on a vector $y\in\mathbb{R}^N$, δ naturally returns a diagonal matrix;

$$\delta(y) \in \mathbb{S}^N \tag{1505}$$

Operating recursively on a vector $\Lambda\in\mathbb{R}^N$ or diagonal matrix $\Lambda\in\mathbb{S}^N$, $\delta\big(\delta(\Lambda)\big)$ returns Λ itself;

$$\delta^2(\Lambda) \equiv \delta\big(\delta(\Lambda)\big) \triangleq \Lambda \tag{1506}$$

Defined in this manner, [243, §3.10, §9.5-1][A.1] main-diagonal linear operator δ is *selfadjoint*; *videlicet*, (§2.2)

$$\delta(A)^{\mathrm T} y = \langle \delta(A)\,,\, y\rangle = \langle A\,,\, \delta(y)\rangle = \operatorname{tr}\big(A^{\mathrm T}\delta(y)\big) \tag{1507}$$

A.1.1 Identities

This δ notation is efficient and unambiguous as illustrated in the following examples where: $A\circ B$ denotes Hadamard product [218] [174, §1.1.4] of matrices of like size, $\otimes$ Kronecker product [182] (§D.1.2.1), y a vector, X a matrix, e_i the i^{th} member of the standard basis for $\mathbb{R}^n$, $\mathbb{S}_h^N$ the symmetric hollow subspace, $\sigma(A)$ a vector of (nonincreasingly) ordered singular values of matrix A, and $\lambda(A)$ a vector of nonincreasingly ordered eigenvalues:

[A.1] Linear operator $T:\mathbb{R}^{m\times n}\rightarrow\mathbb{R}^{M\times N}$ is selfadjoint when, $\forall\, X_1\,,\, X_2\in\mathbb{R}^{m\times n}$

$$\langle T(X_1)\,,\, X_2\rangle = \langle X_1\,,\, T(X_2)\rangle$$

1. $\delta(A) = \delta(A^{\mathrm T})$
2. $\operatorname{tr}(A) = \operatorname{tr}(A^{\mathrm T}) = \delta(A)^{\mathrm T}\mathbf{1} = \langle I\,,A\rangle$
3. $\delta(cA) = c\,\delta(A)$ $\qquad c\in\mathbb{R}$
4. $\operatorname{tr}(cA) = c\operatorname{tr}(A) = c\,\mathbf{1}^{\mathrm T}\lambda(A)$ $\qquad c\in\mathbb{R}$
5. $\operatorname{vec}(cA) = c\operatorname{vec}(A)$ $\qquad c\in\mathbb{R}$
6. $A\circ cB = cA\circ B$ $\qquad c\in\mathbb{R}$
7. $A\otimes cB = cA\otimes B$ $\qquad c\in\mathbb{R}$
8. $\delta(A+B) = \delta(A)+\delta(B)$
9. $\operatorname{tr}(A+B) = \operatorname{tr}(A)+\operatorname{tr}(B)$
10. $\operatorname{vec}(A+B) = \operatorname{vec}(A)+\operatorname{vec}(B)$
11. $(A+B)\circ C = A\circ C+B\circ C$
    $A\circ(B+C) = A\circ B+A\circ C$
12. $(A+B)\otimes C = A\otimes C+B\otimes C$
    $A\otimes(B+C) = A\otimes B+A\otimes C$
13. $\operatorname{sgn}(c)\,\lambda(|c|A) = c\,\lambda(A)$ $\qquad c\in\mathbb{R}$
14. $\operatorname{sgn}(c)\,\sigma(|c|A) = c\,\sigma(A)$ $\qquad c\in\mathbb{R}$
15. $\operatorname{tr}(c\sqrt{A^{\mathrm T}A}\,) = c\operatorname{tr}\sqrt{A^{\mathrm T}A} = c\,\mathbf{1}^{\mathrm T}\sigma(A)$ $\qquad c\in\mathbb{R}$
16. $\pi\boldsymbol{(}\delta(A)\boldsymbol{)} = \lambda(I\circ A)$ where π is the presorting function.
17. $\delta(AB) = (A\circ B^{\mathrm T})\mathbf{1} = (B^{\mathrm T}\circ A)\mathbf{1}$
18. $\delta(AB)^{\mathrm T} = \mathbf{1}^{\mathrm T}(A^{\mathrm T}\circ B) = \mathbf{1}^{\mathrm T}(B\circ A^{\mathrm T})$
19. $\delta(uv^{\mathrm T}) = \begin{bmatrix} u_1v_1 \\ \vdots \\ u_N v_N \end{bmatrix} = u\circ v\,, \quad u,v\in\mathbb{R}^N$
20. $\operatorname{tr}(A^{\mathrm T}B) = \operatorname{tr}(AB^{\mathrm T}) = \operatorname{tr}(BA^{\mathrm T}) = \operatorname{tr}(B^{\mathrm T}A)$
    $= \mathbf{1}^{\mathrm T}(A\circ B)\mathbf{1} = \mathbf{1}^{\mathrm T}\delta(AB^{\mathrm T}) = \delta(A^{\mathrm T}B)^{\mathrm T}\mathbf{1} = \delta(BA^{\mathrm T})^{\mathrm T}\mathbf{1} = \delta(B^{\mathrm T}A)^{\mathrm T}\mathbf{1}$
21. $D=[d_{ij}]\in\mathbb{S}_h^N\,,\quad H=[h_{ij}]\in\mathbb{S}_h^N\,,\quad V=I-\frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm T}\in\mathbb{S}^N$ (*confer* §B.4.2 *no.*20)
    $N\operatorname{tr}(-V(D\circ H)V) = \operatorname{tr}(D^{\mathrm T}H) = \mathbf{1}^{\mathrm T}(D\circ H)\mathbf{1} = \operatorname{tr}\boldsymbol{(}\mathbf{1}\mathbf{1}^{\mathrm T}(D\circ H)\boldsymbol{)} = \sum_{i,j} d_{ij}\,h_{ij}$
22. $\operatorname{tr}(\Lambda A) = \delta(\Lambda)^{\mathrm T}\delta(A)\,,\quad \delta^2(\Lambda)\triangleq\Lambda\in\mathbb{S}^N$
23. $y^{\mathrm T}B\,\delta(A) = \operatorname{tr}\big(B\,\delta(A)y^{\mathrm T}\big) = \operatorname{tr}\big(\delta(B^{\mathrm T}y)A\big) = \operatorname{tr}\big(A\,\delta(B^{\mathrm T}y)\big)$
    $= \delta(A)^{\mathrm T}B^{\mathrm T}y = \operatorname{tr}\big(y\,\delta(A)^{\mathrm T}B^{\mathrm T}\big) = \operatorname{tr}\big(A^{\mathrm T}\delta(B^{\mathrm T}y)\big) = \operatorname{tr}\big(\delta(B^{\mathrm T}y)A^{\mathrm T}\big)$

24. $\delta^2(A^{\mathrm T}A) = \sum_i e_i e_i^{\mathrm T} A^{\mathrm T} A e_i e_i^{\mathrm T}$

25. $\delta\big(\delta(A)\mathbf{1}^{\mathrm T}\big) = \delta\big(\mathbf{1}\,\delta(A)^{\mathrm T}\big) = \delta(A)$

26. $\delta(A\mathbf{1})\mathbf{1} = \delta(A\mathbf{1}\mathbf{1}^{\mathrm T}) = A\mathbf{1}\,, \qquad \delta(y)\mathbf{1} = \delta(y\mathbf{1}^{\mathrm T}) = y$

27. $\delta(I\mathbf{1}) = \delta(\mathbf{1}) = I$

28. $\delta(e_i e_j^{\mathrm T}\mathbf{1}) = \delta(e_i) = e_i e_i^{\mathrm T}$

29. For $\zeta = [\zeta_i] \in \mathbb{R}^k$ and $x = [x_i] \in \mathbb{R}^k$, $\sum_i \zeta_i / x_i = \zeta^{\mathrm T}\delta(x)^{-1}\mathbf{1}$

30. $$\begin{aligned}\operatorname{vec}(A\circ B) &= \operatorname{vec}(A)\circ\operatorname{vec}(B) = \delta(\operatorname{vec}A)\operatorname{vec}(B)\\ &= \operatorname{vec}(B)\circ\operatorname{vec}(A) = \delta(\operatorname{vec}B)\operatorname{vec}(A)\end{aligned} \qquad (42)(1888)$$

31. $\operatorname{vec}(AXB) = (B^{\mathrm T}\otimes A)\operatorname{vec}X$ $\big(\text{not }^{\mathrm H}\big)$

32. $\operatorname{vec}(BXA) = (A^{\mathrm T}\otimes B)\operatorname{vec}X$

33. $$\begin{aligned}\operatorname{tr}(AXBX^{\mathrm T}) &= \operatorname{vec}(X)^{\mathrm T}\operatorname{vec}(AXB) = \operatorname{vec}(X)^{\mathrm T}(B^{\mathrm T}\otimes A)\operatorname{vec}X \quad [182]\\ &= \delta\big(\operatorname{vec}(X)\operatorname{vec}(X)^{\mathrm T}(B^{\mathrm T}\otimes A)\big)^{\mathrm T}\mathbf{1}\end{aligned}$$

34. $$\begin{aligned}\operatorname{tr}(AX^{\mathrm T}BX) &= \operatorname{vec}(X)^{\mathrm T}\operatorname{vec}(BXA) = \operatorname{vec}(X)^{\mathrm T}(A^{\mathrm T}\otimes B)\operatorname{vec}X\\ &= \delta\big(\operatorname{vec}(X)\operatorname{vec}(X)^{\mathrm T}(A^{\mathrm T}\otimes B)\big)^{\mathrm T}\mathbf{1}\end{aligned}$$

35. For any permutation matrix Ξ and dimensionally compatible vector y or matrix A

$$\delta(\Xi y) = \Xi\,\delta(y)\,\Xi^{\mathrm T} \tag{1508}$$
$$\delta(\Xi A \Xi^{\mathrm T}) = \Xi\,\delta(A) \tag{1509}$$

So given any permutation matrix Ξ and any dimensionally compatible matrix B, for example,

$$\delta^2(B) = \Xi\,\delta^2(\Xi^{\mathrm T}B\,\Xi)\Xi^{\mathrm T} \tag{1510}$$

36. $A\otimes 1 = 1\otimes A = A$

37. $A\otimes(B\otimes C) = (A\otimes B)\otimes C$

38. $(A\otimes B)(C\otimes D) = AC\otimes BD$

39. For A a vector, $(A\otimes B) = (A\otimes I)B$

40. For B a row vector, $(A\otimes B) = A(I\otimes B)$

41. $(A\otimes B)^{\mathrm T} = A^{\mathrm T}\otimes B^{\mathrm T}$

42. $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$

43. $\operatorname{tr}(A\otimes B) = \operatorname{tr}A\,\operatorname{tr}B$

44. For $A\in\mathbb{R}^{m\times m}$, $B\in\mathbb{R}^{n\times n}$, $\det(A\otimes B) = \det^n(A)\det^m(B)$

45. There exist permutation matrices Ξ_1 and Ξ_2 such that [182, p.28]

$$A \otimes B = \Xi_1 (B \otimes A) \Xi_2 \tag{1511}$$

46. For eigenvalues $\lambda(A) \in \mathbb{C}^n$ and eigenvectors $v(A) \in \mathbb{C}^{n\times n}$ such that $A = v\delta(\lambda)v^{-1} \in \mathbb{R}^{n\times n}$

$$\lambda(A \otimes B) = \lambda(A) \otimes \lambda(B)\ , \quad v(A \otimes B) = v(A) \otimes v(B) \tag{1512}$$

47. Given *analytic function* [86] $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, then $f(I \otimes A) = I \otimes f(A)$ and $f(A \otimes I) = f(A) \otimes I$ [182, p.28]

A.1.2 Majorization

A.1.2.0.1 Theorem. (Schur) *Majorization.* [432, §7.4] [218, §4.3] [219, §5.5]
Let $\lambda \in \mathbb{R}^N$ denote a given vector of eigenvalues and let $\delta \in \mathbb{R}^N$ denote a given vector of main diagonal entries, both arranged in nonincreasing order. Then

$$\exists\, A \in \mathbb{S}^N \ni \lambda(A) = \lambda \ \textbf{ and } \ \delta(A) = \delta \quad \Leftarrow \quad \lambda - \delta \in \mathcal{K}^*_{\lambda\delta} \tag{1513}$$

and conversely

$$A \in \mathbb{S}^N \ \Rightarrow \ \lambda(A) - \delta(A) \in \mathcal{K}^*_{\lambda\delta} \tag{1514}$$

The difference belongs to the pointed polyhedral cone of majorization (not a full-dimensional cone, *confer* (312))

$$\mathcal{K}^*_{\lambda\delta} \triangleq \mathcal{K}^*_{\mathcal{M}+} \cap \{\zeta \mathbf{1} \mid \zeta \in \mathbb{R}\}^* \tag{1515}$$

where $\mathcal{K}^*_{\mathcal{M}+}$ is the dual monotone nonnegative cone (434), and where the dual of the line is a hyperplane; $\partial\mathcal{H} = \{\zeta\mathbf{1} \mid \zeta \in \mathbb{R}\}^* = \mathbf{1}^\perp$. ⋄

Majorization cone $\mathcal{K}^*_{\lambda\delta}$ is naturally consequent to the definition of majorization; *id est*, vector $y \in \mathbb{R}^N$ *majorizes* vector x if and only if

$$\sum_{i=1}^{k} x_i \le \sum_{i=1}^{k} y_i \quad \forall\, 1 \le k \le N \tag{1516}$$

and

$$\mathbf{1}^{\mathrm T} x = \mathbf{1}^{\mathrm T} y \tag{1517}$$

Under these circumstances, rather, vector x is majorized by vector y.

In the particular circumstance $\delta(A) = \mathbf{0}$ we get:

A.1.2.0.2 Corollary. *Symmetric hollow majorization.*
Let $\lambda \in \mathbb{R}^N$ denote a given vector of eigenvalues arranged in nonincreasing order. Then

$$\exists\, A \in \mathbb{S}^N_h \ni \lambda(A) = \lambda \quad \Leftarrow \quad \lambda \in \mathcal{K}^*_{\lambda\delta} \tag{1518}$$

and conversely

$$A \in \mathbb{S}^N_h \ \Rightarrow \ \lambda(A) \in \mathcal{K}^*_{\lambda\delta} \tag{1519}$$

where $\mathcal{K}^*_{\lambda\delta}$ is defined in (1515). ⋄

A.2 Semidefiniteness: domain of test

The most fundamental necessary, sufficient, and definitive test for positive semidefiniteness of matrix $A\in\mathbb{R}^{n\times n}$ is: [219, §1]

$$x^{\mathrm{T}}A\,x \geq 0 \;\text{ for each and every } x\in\mathbb{R}^n \text{ such that } \|x\|=1 \tag{1520}$$

Traditionally, authors demand evaluation over broader domain; namely, over all $x\in\mathbb{R}^n$ which is sufficient but unnecessary. Indeed, that standard textbook requirement is far over-reaching because if $x^{\mathrm{T}}A\,x$ is nonnegative for particular $x=x_{\mathrm{p}}$, then it is nonnegative for any αx_{p} where $\alpha\in\mathbb{R}$. Thus, only normalized x in $\mathbb{R}^n$ need be evaluated.

Many authors add the further requirement that the domain be complex; the broadest domain. By so doing, only *Hermitian matrices* ($A^{\mathrm{H}}=A$ where superscript $^{\mathrm{H}}$ denotes conjugate transpose)[A.2] are admitted to the set of positive semidefinite matrices (1523); an unnecessary prohibitive condition.

A.2.1 Symmetry *versus* semidefiniteness

We call (1520) *the most fundamental test* of positive semidefiniteness. Yet some authors instead say, for real A and complex domain $\{x\in\mathbb{C}^n\}$, the complex test $x^{\mathrm{H}}Ax\geq 0$ is most fundamental. That complex broadening of the domain of test causes nonsymmetric real matrices to be excluded from the set of positive semidefinite matrices. Yet admitting nonsymmetric real matrices or not is a matter of preference[A.3] unless that complex test is adopted, as we shall now explain.

Any real square matrix A has a representation in terms of its symmetric and antisymmetric parts; *id est*,

$$A=\frac{(A+A^{\mathrm{T}})}{2}+\frac{(A-A^{\mathrm{T}})}{2} \tag{53}$$

Because, for all real A , the antisymmetric part vanishes under real test,

$$x^{\mathrm{T}}\frac{(A-A^{\mathrm{T}})}{2}\,x=0 \tag{1521}$$

only the symmetric part of A , $(A+A^{\mathrm{T}})/2$, has a role determining positive semidefiniteness. Hence the oft-made presumption that only symmetric matrices may be positive semidefinite is, of course, erroneous under (1520). Because eigenvalue-signs of a symmetric matrix translate unequivocally to its semidefiniteness, the eigenvalues that determine semidefiniteness are always those of the *symmetrized* matrix. (§A.3) For that reason, and because symmetric (or Hermitian) matrices must have real eigenvalues, the convention adopted in the literature is that semidefinite matrices are synonymous with symmetric semidefinite matrices. Certainly misleading under (1520), that presumption is typically bolstered with compelling examples from the physical sciences where symmetric matrices occur within the mathematical exposition of natural phenomena.[A.4] [152, §52]

[A.2] Hermitian symmetry is the complex analogue; the real part of a Hermitian matrix is symmetric while its imaginary part is antisymmetric. A Hermitian matrix has real eigenvalues and real main diagonal.

[A.3] Golub & Van Loan [174, §4.2.2], for example, admit nonsymmetric real matrices.

[A.4] Symmetric matrices are certainly pervasive in the our chosen subject as well.

Perhaps a better explanation of this pervasive presumption of symmetry comes from Horn & Johnson [218, §7.1] whose perspective[A.5] is the complex matrix, thus necessitating the complex domain of test throughout their treatise. They explain, if $A \in \mathbb{C}^{n\times n}$

> *. . . and if $x^{\rm H}Ax$ is real for all $x \in \mathbb{C}^n$, then A is Hermitian. Thus, the assumption that A is Hermitian is not necessary in the definition of positive definiteness. It is customary, however.*

Their comment is best explained by noting, the real part of $x^{\rm H}Ax$ comes from the Hermitian part $(A + A^{\rm H})/2$ of A;

$$\mathrm{re}(x^{\rm H}Ax) = x^{\rm H}\,\frac{A + A^{\rm H}}{2}\,x \tag{1522}$$

rather,

$$x^{\rm H}Ax \in \mathbb{R} \;\Leftrightarrow\; A^{\rm H} = A \tag{1523}$$

because the imaginary part of $x^{\rm H}Ax$ comes from the anti-Hermitian part $(A - A^{\rm H})/2$;

$$\mathrm{im}(x^{\rm H}Ax) = x^{\rm H}\,\frac{A - A^{\rm H}}{2}\,x \tag{1524}$$

that vanishes for nonzero x if and only if $A = A^{\rm H}$. So the Hermitian symmetry assumption is unnecessary, according to Horn & Johnson, not because nonHermitian matrices could be regarded positive semidefinite, rather because nonHermitian (includes nonsymmetric real) matrices are not comparable on the real line under $x^{\rm H}Ax$. Yet that complex edifice is dismantled in the test of real matrices (1520) because the domain of test is no longer necessarily complex; meaning, $x^{\rm T}Ax$ will certainly always be real, regardless of symmetry, and so real A will always be comparable.

In summary, if we limit the domain of test to all x in $\mathbb{R}^n$ as in (1520), then nonsymmetric real matrices are admitted to the realm of semidefinite matrices because they become comparable on the real line. One important exception occurs for rank-one matrices $\Psi = uv^{\rm T}$ where u and v are real vectors: Ψ is positive semidefinite if and only if $\Psi = uu^{\rm T}$. (§A.3.1.0.7)

We might choose to expand the domain of test to all x in $\mathbb{C}^n$ so that only symmetric matrices would be comparable. The alternative to expanding domain of test is to assume all matrices of interest to be symmetric; that is commonly done, hence the synonymous relationship with semidefinite matrices.

A.2.1.0.1 Example. *Nonsymmetric positive definite product.*
Horn & Johnson assert and Zhang agrees:

> *If $A, B \in \mathbb{C}^{n\times n}$ are positive definite, then we know that the product AB is positive definite if and only if AB is Hermitian.* [218, §7.6 prob.10] [432, §6.2, §3.2]

[A.5] A totally complex perspective is not necessarily more advantageous. The positive semidefinite cone, for example, is not selfdual (§2.13.5) in the ambient space of Hermitian matrices. [211, §II]

Implicitly in their statement, A and B are assumed individually Hermitian and the domain of test is assumed complex.

We prove that assertion to be false for real matrices under (1520) that adopts a real domain of test.

$$A^{\mathrm{T}} = A = \begin{bmatrix} 3 & 0 & -1 & 0 \\ 0 & 5 & 1 & 0 \\ -1 & 1 & 4 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}, \qquad \lambda(A) = \begin{bmatrix} 5.9 \\ 4.5 \\ 3.4 \\ 2.0 \end{bmatrix} \tag{1525}$$

$$B^{\mathrm{T}} = B = \begin{bmatrix} 4 & 4 & -1 & -1 \\ 4 & 5 & 0 & 0 \\ -1 & 0 & 5 & 1 \\ -1 & 0 & 1 & 4 \end{bmatrix}, \qquad \lambda(B) = \begin{bmatrix} 8.8 \\ 5.5 \\ 3.3 \\ 0.24 \end{bmatrix} \tag{1526}$$

$$(AB)^{\mathrm{T}} \neq AB = \begin{bmatrix} 13 & 12 & -8 & -4 \\ 19 & 25 & 5 & 1 \\ -5 & 1 & 22 & 9 \\ -5 & 0 & 9 & 17 \end{bmatrix}, \qquad \lambda(AB) = \begin{bmatrix} 36. \\ 29. \\ 10. \\ 0.72 \end{bmatrix} \tag{1527}$$

$$\frac{1}{2}(AB+(AB)^{\mathrm{T}}) = \begin{bmatrix} 13 & 15.5 & -6.5 & -4.5 \\ 15.5 & 25 & 3 & 0.5 \\ -6.5 & 3 & 22 & 9 \\ -4.5 & 0.5 & 9 & 17 \end{bmatrix}, \quad \lambda\left(\frac{1}{2}(AB + (AB)^{\mathrm{T}})\right) = \begin{bmatrix} 36. \\ 30. \\ 10. \\ 0.014 \end{bmatrix} \tag{1528}$$

Whenever $A \in \mathbb{S}^n_+$ and $B \in \mathbb{S}^n_+$, then $\lambda(AB) = \lambda(\sqrt{A}B\sqrt{A})$ will always be a nonnegative vector by (1552) and Corollary A.3.1.0.5. Yet positive definiteness of product AB is certified instead by the nonnegative eigenvalues $\lambda\left(\frac{1}{2}(AB + (AB)^{\mathrm{T}})\right)$ in (1528) (§A.3.1.0.1) despite the fact that AB is not symmetric.[A.6] ♦

Horn & Johnson and Zhang resolve this anomaly by choosing to exclude nonsymmetric matrices and products; they do so by expanding the domain of test to $\mathbb{C}^n$. □

A.3 Proper statements of positive semidefiniteness

Unlike Horn & Johnson and Zhang, we never adopt a complex domain of test with real matrices. So motivated is our consideration of proper statements of positive semidefiniteness under real domain of test. This restriction, ironically, complicates the facts when compared to corresponding statements for the complex case (found elsewhere [218] [432]).

We state several fundamental facts regarding positive semidefiniteness of real matrix A and the product AB and sum $A+B$ of real matrices under fundamental real test (1520); a few require proof as they depart from the standard texts, while those remaining are well established or obvious.

[A.6] It is a little more difficult to find a counterexample in $\mathbb{R}^{2\times 2}$ or $\mathbb{R}^{3\times 3}$; which may have served to advance any confusion.

A.3.0.0.1 Theorem. *Positive (semi)definite matrix.*
$A\in\mathbb{S}^M$ is positive semidefinite if and only if for each and every vector $x\in\mathbb{R}^M$ of unit norm, $\|x\|=1$,[A.7] we have $x^{\mathrm T}A\,x\geq 0$ (1520);

$$A\succeq 0 \;\Leftrightarrow\; \operatorname{tr}(xx^{\mathrm T}A) \,=\, x^{\mathrm T}A\,x\geq 0 \;\;\forall\, xx^{\mathrm T} \tag{1529}$$

Matrix $A\in\mathbb{S}^M$ is positive definite if and only if for each and every $\|x\|=1$ we have $x^{\mathrm T}A\,x>0$;

$$A\succ 0 \;\Leftrightarrow\; \operatorname{tr}(xx^{\mathrm T}A) \,=\, x^{\mathrm T}A\,x> 0 \;\;\forall\, xx^{\mathrm T},\;\; xx^{\mathrm T}\neq \mathbf{0} \tag{1530}$$

⋄

Proof. Statements (1529) and (1530) are each a particular instance of dual generalized inequalities (§2.13.2) with respect to the positive semidefinite cone; *videlicet*, [379]

$$\begin{array}{l} A\succeq 0 \;\Leftrightarrow\; \langle xx^{\mathrm T},\,A\rangle\geq 0 \;\;\forall\, xx^{\mathrm T}(\succeq 0)\\ A\succ 0 \;\Leftrightarrow\; \langle xx^{\mathrm T},\,A\rangle> 0 \;\;\forall\, xx^{\mathrm T}(\succeq 0)\,,\;\; xx^{\mathrm T}\neq \mathbf{0} \end{array} \tag{1531}$$

This says: positive semidefinite matrix A must belong to the normal side of every hyperplane whose normal is an extreme direction of the positive semidefinite cone. Relations (1529) and (1530) remain true when $xx^{\mathrm T}$ is replaced with "for each and every" positive semidefinite matrix $X\in\mathbb{S}_+^M$ (§2.13.5) of unit norm, $\|X\|=1$, as in

$$\begin{array}{l} A\succeq 0 \;\Leftrightarrow\; \operatorname{tr}(XA)\geq 0 \;\;\forall\, X\in\mathbb{S}_+^M\\ A\succ 0 \;\Leftrightarrow\; \operatorname{tr}(XA)> 0 \;\;\forall\, X\in\mathbb{S}_+^M\,,\;\; X\neq \mathbf{0} \end{array} \tag{1532}$$

But that condition is more than what is necessary. By the *discretized membership theorem* in §2.13.4.2.1, the extreme directions $xx^{\mathrm T}$ of the positive semidefinite cone constitute a minimal set of generators necessary and sufficient for discretization of dual generalized inequalities (1532) certifying membership to that cone. ◆

A.3.1 Semidefiniteness, eigenvalues, nonsymmetric

When $A\in\mathbb{R}^{n\times n}$, let $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)\in\mathbb{R}^n$ denote eigenvalues of the symmetrized matrix[A.8] arranged in nonincreasing order.

- By positive semidefiniteness of $A\in\mathbb{R}^{n\times n}$ we mean,[A.9]
  [288, §1.3.1] (*confer* §A.3.1.0.1)

$$x^{\mathrm T}A\,x\geq 0\;\;\forall\, x\in\mathbb{R}^n \;\Leftrightarrow\; A+A^{\mathrm T}\succeq 0 \;\Leftrightarrow\; \lambda(A+A^{\mathrm T})\succeq 0 \tag{1533}$$

[A.7] The traditional condition requiring all $x\in\mathbb{R}^M$ for defining positive (semi)definiteness is actually more than what is necessary. The set of norm-1 vectors is necessary and sufficient to establish positive semidefiniteness; actually, any particular norm and any nonzero norm-constant will work.

[A.8] The symmetrization of A is $(A+A^{\mathrm T})/2$. $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)=\lambda(A+A^{\mathrm T})/2$.

[A.9] Strang agrees [348, p.334] it is not $\lambda(A)$ that requires observation. Yet he is mistaken by proposing the Hermitian part alone $x^{\mathrm H}(A+A^{\mathrm H})x$ be tested, because the anti-Hermitian part does not vanish under complex test unless A is Hermitian. (1524)

- (§2.9.0.1)

$$A \succeq 0 \;\Rightarrow\; A^{\mathrm T} = A \tag{1534}$$

$$A \succeq B \;\Leftrightarrow\; A - B \succeq 0 \;\nRightarrow\; A \succeq 0 \;\text{ or }\; B \succeq 0 \tag{1535}$$

$$x^{\mathrm T} A\, x \geq 0 \;\;\forall\, x \;\nRightarrow\; A^{\mathrm T} = A \tag{1536}$$

- Matrix symmetry is not intrinsic to positive semidefiniteness;

$$A^{\mathrm T} = A\,,\quad \lambda(A) \succeq 0 \;\Rightarrow\; x^{\mathrm T} A\, x \geq 0 \;\;\forall\, x \tag{1537}$$

$$\lambda(A) \succeq 0 \;\Leftarrow\; A^{\mathrm T} = A\,,\quad x^{\mathrm T} A\, x \geq 0 \;\;\forall\, x \tag{1538}$$

- If $A^{\mathrm T} = A$ then

$$\lambda(A) \succeq 0 \;\Leftrightarrow\; A \succeq 0 \tag{1539}$$

meaning, matrix A belongs to the positive semidefinite cone in the subspace of symmetric matrices if and only if its eigenvalues belong to the nonnegative orthant.

$$\langle A\;,\;A\rangle = \langle \lambda(A)\,,\,\lambda(A)\rangle \qquad (45)$$

- For $\mu \in \mathbb{R}$, $A \in \mathbb{R}^{n\times n}$, and vector $\lambda(A) \in \mathbb{C}^n$ holding the ordered eigenvalues of A

$$\lambda(\mu I + A) = \mu\mathbf{1} + \lambda(A) \tag{1540}$$

Proof. $A = MJM^{-1}$ and $\mu I + A = M(\mu I + J)M^{-1}$ where J is the *Jordan form* for A; [348, §5.6, App.B] *id est*, $\delta(J) = \lambda(A)$, so $\lambda(\mu I + A) = \delta(\mu I + J)$ because $\mu I + J$ is also a Jordan form. ♦

By similar reasoning,

$$\lambda(I + \mu A) = \mathbf{1} + \lambda(\mu A) \tag{1541}$$

For vector $\sigma(A)$ holding the singular values of any matrix A

$$\sigma(I + \mu A^{\mathrm T}A) = \pi\big(|\mathbf{1} + \mu\sigma(A^{\mathrm T}A)|\big) \tag{1542}$$

$$\sigma(\mu I + A^{\mathrm T}A) = \pi\big(|\mu\mathbf{1} + \sigma(A^{\mathrm T}A)|\big) \tag{1543}$$

where π is the nonlinear permutation-operator sorting its vector argument into nonincreasing order.

- For $A \in \mathbb{S}^M$ and each and every $\|w\| = 1$ [218, §7.7 prob.9]

$$w^{\mathrm T}Aw \leq \mu \;\Leftrightarrow\; A \preceq \mu I \;\Leftrightarrow\; \lambda(A) \preceq \mu\mathbf{1} \tag{1544}$$

- [218, §2.5.4] (*confer* (44))

$$A \text{ is normal matrix } \;\Leftrightarrow\; \|A\|_{\mathrm F}^2 = \lambda(A)^{\mathrm T}\lambda(A) \tag{1545}$$

- For $A\in\mathbb{R}^{m\times n}$
$$A^{\mathrm T}A\succeq 0\,,\qquad AA^{\mathrm T}\succeq 0 \tag{1546}$$
because, for dimensionally compatible vector x,
$x^{\mathrm T}A^{\mathrm T}Ax=\|Ax\|_2^2$, $x^{\mathrm T}AA^{\mathrm T}x=\|A^{\mathrm T}x\|_2^2$.

- For $A\in\mathbb{R}^{n\times n}$ and $c\in\mathbb{R}$
$$\operatorname{tr}(cA)\;=\;c\operatorname{tr}(A)\;=\;c\,\mathbf{1}^{\mathrm T}\lambda(A)\qquad(\S\text{A.1.1 } no.4)$$
For m a nonnegative integer, (1969)
$$\det(A^m)=\prod_{i=1}^{n}\lambda(A)_i^m \tag{1547}$$
$$\operatorname{tr}(A^m)=\sum_{i=1}^{n}\lambda(A)_i^m \tag{1548}$$

- For A diagonalizable (§A.5), $A=S\Lambda S^{-1}$, (*confer* [348, p.255])
$$\operatorname{rank}A=\operatorname{rank}\delta(\lambda(A))=\operatorname{rank}\Lambda \tag{1549}$$
meaning, rank is equal to the number of nonzero eigenvalues in vector
$$\lambda(A)\triangleq\delta(\Lambda) \tag{1550}$$
by the 0 *eigenvalues theorem* (§A.7.3.0.1).

- (Ky Fan) For $A\,,B\in\mathbb{S}^n$ [56, §1.2] (*confer* (1841))
$$\operatorname{tr}(AB)\leq\lambda(A)^{\mathrm T}\lambda(B)\qquad(1827)$$
with equality (Theobald) when A and B are simultaneously diagonalizable [218] with the same ordering of eigenvalues.

- For $A\in\mathbb{R}^{m\times n}$ and $B\in\mathbb{R}^{n\times m}$
$$\operatorname{tr}(AB)=\operatorname{tr}(BA) \tag{1551}$$
and η eigenvalues of the product and commuted product are identical, including their multiplicity; [218, §1.3.20] *id est*,
$$\lambda(AB)_{1:\eta}=\lambda(BA)_{1:\eta}\;,\qquad \eta\triangleq\min\{m\,,n\} \tag{1552}$$
Any eigenvalues remaining are zero. By the 0 *eigenvalues theorem* (§A.7.3.0.1),
$$\operatorname{rank}(AB)=\operatorname{rank}(BA)\,,\qquad AB\text{ and }BA\text{ diagonalizable} \tag{1553}$$

- For any compatible matrices $A\,,B$ [218, §0.4]
$$\min\{\operatorname{rank}A\,,\operatorname{rank}B\}\geq\operatorname{rank}(AB) \tag{1554}$$

- For $A, B \in \mathbb{S}^n_+$ (*confer* (257))

$$\operatorname{rank} A + \operatorname{rank} B \geq \operatorname{rank}(A+B) \geq \min\{\operatorname{rank} A\,,\ \operatorname{rank} B\} \geq \operatorname{rank}(AB) \tag{1555}$$

- For linearly independent matrices $A, B \in \mathbb{S}^n_+$ (§2.1.2, $\mathcal{R}(A) \cap \mathcal{R}(B) = \mathbf{0}$, $\mathcal{R}(A^{\mathrm{T}}) \cap \mathcal{R}(B^{\mathrm{T}}) = \mathbf{0}$, §B.1.1),

$$\operatorname{rank} A + \operatorname{rank} B = \operatorname{rank}(A+B) > \min\{\operatorname{rank} A\,,\ \operatorname{rank} B\} \geq \operatorname{rank}(AB) \tag{1556}$$

- Because $\mathcal{R}(A^{\mathrm{T}}A) = \mathcal{R}(A^{\mathrm{T}})$ and $\mathcal{R}(AA^{\mathrm{T}}) = \mathcal{R}(A)$ (p.548), for any $A \in \mathbb{R}^{m\times n}$

$$\operatorname{rank}(AA^{\mathrm{T}}) = \operatorname{rank}(A^{\mathrm{T}}A) = \operatorname{rank} A = \operatorname{rank} A^{\mathrm{T}} \tag{1557}$$

- For $A \in \mathbb{R}^{m\times n}$ having no nullspace, and for any $B \in \mathbb{R}^{n\times k}$

$$\operatorname{rank}(AB) = \operatorname{rank}(B) \tag{1558}$$

Proof. For any compatible matrix C, $\mathcal{N}(CAB) \supseteq \mathcal{N}(AB) \supseteq \mathcal{N}(B)$ is obvious. By assumption $\exists A^{\dagger} \ni A^{\dagger}A = I$. Let $C = A^{\dagger}$, then $\mathcal{N}(AB) = \mathcal{N}(B)$ and the stated result follows by conservation of dimension (1674). ♦

- For $A \in \mathbb{S}^n$ and any nonsingular matrix Y

$$\operatorname{inertia}(A) = \operatorname{inertia}(YAY^{\mathrm{T}}) \tag{1559}$$

`a.k.a` *Sylvester's law of inertia.* (1605) [120, §2.4.3]

- For $A, B \in \mathbb{R}^{n\times n}$ square, [218, §0.3.5]

$$\det(AB) = \det(BA) \tag{1560}$$

$$\det(AB) = \det A \det B \tag{1561}$$

Yet for $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$ [80, p.72]

$$\det(I + AB) = \det(I + BA) \tag{1562}$$

- For $A, B \in \mathbb{S}^n$, product AB is symmetric iff AB is commutative;

$$(AB)^{\mathrm{T}} = AB \;\Leftrightarrow\; AB = BA \tag{1563}$$

Proof. (⇒) Suppose $AB = (AB)^{\mathrm{T}}$.
$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}} = BA$. $AB = (AB)^{\mathrm{T}} \Rightarrow AB = BA$.
(⇐) Suppose $AB = BA$.
$BA = B^{\mathrm{T}}A^{\mathrm{T}} = (AB)^{\mathrm{T}}$. $AB = BA \Rightarrow AB = (AB)^{\mathrm{T}}$. ♦

Commutativity alone is insufficient for product symmetry. [348, p.26] Matrix symmetry alone is insufficient for product symmetry.

- Diagonalizable matrices $A\,,B\in\mathbb{R}^{n\times n}$ commute if and only if they are simultaneously diagonalizable. [218, §1.3.12] A product of diagonal matrices is always commutative.

- For $A,B\in\mathbb{R}^{n\times n}$ and $AB=BA$

$$x^{\mathrm T}A\,x\geq 0\,,\ x^{\mathrm T}B\,x\geq 0\ \forall x\ \Rightarrow\ \lambda(A+A^{\mathrm T})_i\,\lambda(B+B^{\mathrm T})_i\geq 0\ \forall i\ \not\Leftrightarrow\ x^{\mathrm T}AB\,x\geq 0\ \forall x \tag{1564}$$

the negative result arising because of the schism between the product of eigenvalues $\lambda(A+A^{\mathrm T})_i\,\lambda(B+B^{\mathrm T})_i$ and the eigenvalues of the symmetrized matrix product $\lambda\big(AB+(AB)^{\mathrm T}\big)_i$. For example, X^2 is generally not positive semidefinite unless matrix X is symmetric; then (1546) applies. Simply substituting symmetric matrices changes the outcome:

- For $A,B\in\mathbb{S}^n$ and $AB=BA$

$$A\succeq 0\,,\ B\succeq 0\ \Rightarrow\ \lambda(AB)_i=\lambda(A)_i\,\lambda(B)_i\geq 0\ \ \forall i\ \ \Leftrightarrow\ AB\succeq 0 \tag{1565}$$

Positive semidefiniteness of commutative A and B is sufficient but not necessary for positive semidefiniteness of product AB .

Proof. Because all symmetric matrices are diagonalizable, (§A.5.1) [348, §5.6] we have $A=S\Lambda S^{\mathrm T}$ and $B=T\Delta T^{\mathrm T}$, where Λ and Δ are real diagonal matrices while S and T are orthogonal matrices. Because $(AB)^{\mathrm T}=AB$, then T must equal S , [218, §1.3] and the eigenvalues of A are ordered in the same way as those of B ; *id est*, $\lambda(A)_i=\delta(\Lambda)_i$ and $\lambda(B)_i=\delta(\Delta)_i$ correspond to the same eigenvector.
($\Rightarrow$) Assume $\lambda(A)_i\,\lambda(B)_i\geq 0$ for $i=1\ldots n$. $AB=S\Lambda\Delta S^{\mathrm T}$ is symmetric and has nonnegative eigenvalues contained in diagonal matrix $\Lambda\Delta$ by assumption; hence positive semidefinite by (1533). Now assume $A,B\succeq 0$. That, of course, implies $\lambda(A)_i\,\lambda(B)_i\geq 0$ for all i because all the individual eigenvalues are nonnegative.
($\Leftarrow$) Suppose $AB=S\Lambda\Delta S^{\mathrm T}\succeq 0$. Then $\Lambda\Delta\succeq 0$ by (1533), and so all products $\lambda(A)_i\,\lambda(B)_i$ must be nonnegative; meaning, $\mathrm{sgn}(\lambda(A))=\mathrm{sgn}(\lambda(B))$. We may, therefore, conclude nothing about the semidefiniteness of A and B . ♦

- For $A,B\in\mathbb{S}^n$ and $A\succeq 0\,,\ B\succeq 0$ (Example A.2.1.0.1)

$$AB=BA\ \Rightarrow\ \lambda(AB)_i=\lambda(A)_i\,\lambda(B)_i\geq 0\ \ \forall i\ \ \Rightarrow\ AB\succeq 0 \tag{1566}$$

$$AB=BA\ \Rightarrow\ \lambda(AB)_i\geq 0\,,\ \ \lambda(A)_i\,\lambda(B)_i\geq 0\ \ \forall i\ \ \Leftrightarrow\ AB\succeq 0 \tag{1567}$$

- For $A,B\in\mathbb{S}^n$ [218, §7.7 prob.3] [219, §4.2.13, §5.2.1]

$$A\succeq 0\,,\ B\succeq 0\ \Rightarrow\ A\otimes B\succeq 0 \tag{1568}$$
$$A\succeq 0\,,\ B\succeq 0\ \Rightarrow\ A\circ B\succeq 0 \tag{1569}$$
$$A\succ 0\,,\ B\succ 0\ \Rightarrow\ A\otimes B\succ 0 \tag{1570}$$
$$A\succ 0\,,\ B\succ 0\ \Rightarrow\ A\circ B\succ 0 \tag{1571}$$

where Kronecker and Hadamard products are symmetric.

- For $A, B \in \mathbb{S}^n$, (1539) $A \succeq 0 \Leftrightarrow \lambda(A) \succeq 0$ yet

$$A \succeq 0 \;\Rightarrow\; \delta(A) \succeq 0 \tag{1572}$$

$$A \succeq 0 \;\Rightarrow\; \operatorname{tr} A \geq 0 \tag{1573}$$

$$A \succeq 0\,,\; B \succeq 0 \;\Rightarrow\; \operatorname{tr} A \operatorname{tr} B \geq \operatorname{tr}(AB) \geq 0 \tag{1574}$$

[432, §6.2] Because $A \succeq 0\,,\; B \succeq 0 \;\Rightarrow\; \lambda(AB) = \lambda(\sqrt{A}\,B\sqrt{A}) \succeq 0$ by (1552) and Corollary A.3.1.0.5, then we have $\operatorname{tr}(AB) \geq 0$.

$$A \succeq 0 \;\Leftrightarrow\; \operatorname{tr}(AB) \geq 0 \quad \forall\, B \succeq 0 \qquad (377)$$

- For $A, B, C \in \mathbb{S}^n$ (Löwner)

$$\begin{array}{lr} A \preceq B\,,\;\; B \preceq C \;\Rightarrow\; A \preceq C & \text{(transitivity)} \\ A \preceq B \;\Leftrightarrow\; A + C \preceq B + C & \text{(additivity)} \\ A \preceq B\,,\;\; A \succeq B \;\Rightarrow\; A = B & \text{(antisymmetry)} \\ A \preceq A & \text{(reflexivity)} \end{array} \tag{1575}$$

$$\begin{array}{lr} A \preceq B\,,\;\; B \prec C \;\Rightarrow\; A \prec C & \text{(strict transitivity)} \\ A \prec B \;\Leftrightarrow\; A + C \prec B + C & \text{(strict additivity)} \end{array} \tag{1576}$$

- For $A, B \in \mathbb{R}^{n \times n}$

$$x^{\mathrm{T}} A x \geq x^{\mathrm{T}} B x \;\; \forall x \;\Rightarrow\; \operatorname{tr} A \geq \operatorname{tr} B \tag{1577}$$

Proof. $x^{\mathrm{T}} A x \geq x^{\mathrm{T}} B x \;\; \forall x \;\Leftrightarrow\; \lambda\big((A-B) + (A-B)^{\mathrm{T}}\big)/2 \succeq 0 \;\Rightarrow$ $\operatorname{tr}\big(A + A^{\mathrm{T}} - (B + B^{\mathrm{T}})\big)/2 = \operatorname{tr}(A - B) \geq 0$. There is no converse. ♦

- For $A, B \in \mathbb{S}^n$ [432, §6.2] (Theorem A.3.1.0.4)

$$A \succeq B \;\Rightarrow\; \operatorname{tr} A \geq \operatorname{tr} B \tag{1578}$$

$$A \succeq B \;\Rightarrow\; \delta(A) \succeq \delta(B) \tag{1579}$$

There is no converse, and restriction to the positive semidefinite cone does not improve the situation. All-strict versions hold.

$$A \succeq B \succeq 0 \;\Rightarrow\; \operatorname{rank} A \geq \operatorname{rank} B \tag{1580}$$

$$A \succeq B \succeq 0 \;\Rightarrow\; \det A \geq \det B \geq 0 \tag{1581}$$

$$A \succ B \succeq 0 \;\Rightarrow\; \det A > \det B \geq 0 \tag{1582}$$

- For $A, B \in \operatorname{int} \mathbb{S}^n_+$ [35, §4.2] [218, §7.7.4]

$$A \succeq B \;\Leftrightarrow\; A^{-1} \preceq B^{-1}, \qquad A \succ 0 \;\Leftrightarrow\; A^{-1} \succ 0 \tag{1583}$$

- For $A, B \in \mathbb{S}^n$ [432, §6.2]

$$\begin{array}{r} A \succeq B \succeq 0 \;\Rightarrow\; \sqrt{A} \succeq \sqrt{B} \\ A \succeq 0 \;\Leftarrow\; A^{1/2} \succeq 0 \end{array} \tag{1584}$$

- For $A, B \in \mathbb{S}^n$ and $AB = BA$ [432, §6.2 prob.3]

$$A \succeq B \succeq 0 \;\Rightarrow\; A^k \succeq B^k\,, \quad k = 1, 2, \ldots \tag{1585}$$

A.3.1.0.1 Theorem. *Positive semidefinite ordering of eigenvalues.*
For $A, B \in \mathbb{R}^{M\times M}$, place the eigenvalues of each symmetrized matrix into the respective vectors $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)\,,\ \lambda\big(\frac{1}{2}(B+B^{\mathrm T})\big)\in\mathbb{R}^M$. Then [348, §6]

$$x^{\mathrm T}A\,x \geq 0 \quad \forall x \qquad \Leftrightarrow \quad \lambda\big(A+A^{\mathrm T}\big)\succeq 0 \tag{1586}$$

$$x^{\mathrm T}A\,x > 0 \quad \forall x\neq \mathbf{0} \quad \Leftrightarrow \quad \lambda\big(A+A^{\mathrm T}\big)\succ 0 \tag{1587}$$

because $x^{\mathrm T}(A-A^{\mathrm T})x=0$. (1521) Now arrange the entries of $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)$ and $\lambda\big(\frac{1}{2}(B+B^{\mathrm T})\big)$ in nonincreasing order so $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)_1$ holds the largest eigenvalue of symmetrized A while $\lambda\big(\frac{1}{2}(B+B^{\mathrm T})\big)_1$ holds the largest eigenvalue of symmetrized B, and so on. Then [218, §7.7 prob.1 prob.9] for $\kappa\in\mathbb{R}$

$$\begin{array}{rcl} x^{\mathrm T}A\,x \geq x^{\mathrm T}B\,x \quad \forall x & \Rightarrow & \lambda\big(A+A^{\mathrm T}\big)\succeq \lambda\big(B+B^{\mathrm T}\big) \\ x^{\mathrm T}A\,x \geq x^{\mathrm T}I\,x\,\kappa \quad \forall x & \Leftrightarrow & \lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)\succeq \kappa\mathbf{1} \end{array} \tag{1588}$$

Now let $A, B\in\mathbb{S}^M$ have diagonalizations $A=Q\Lambda Q^{\mathrm T}$ and $B=U\Upsilon U^{\mathrm T}$ with $\lambda(A)=\delta(\Lambda)$ and $\lambda(B)=\delta(\Upsilon)$ arranged in nonincreasing order. Then

$$A\succeq B \;\Leftrightarrow\; \lambda(A-B)\succeq 0 \tag{1589}$$

$$A\succeq B \;\Rightarrow\; \lambda(A)\succeq \lambda(B) \tag{1590}$$

$$A\succeq B \;\not\Leftarrow\; \lambda(A)\succeq \lambda(B) \tag{1591}$$

$$S^{\mathrm T}AS\succeq B \;\Leftarrow\; \lambda(A)\succeq \lambda(B) \tag{1592}$$

where $S=QU^{\mathrm T}$. [432, §7.5] ⋄

A.3.1.0.2 Theorem. (Weyl) *Eigenvalues of sum.* [218, §4.3.1]
For $A, B\in\mathbb{R}^{M\times M}$, place the eigenvalues of each symmetrized matrix into the respective vectors $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)\,,\ \lambda\big(\frac{1}{2}(B+B^{\mathrm T})\big)\in\mathbb{R}^M$ in nonincreasing order so $\lambda\big(\frac{1}{2}(A+A^{\mathrm T})\big)_1$ holds the largest eigenvalue of symmetrized A while $\lambda\big(\frac{1}{2}(B+B^{\mathrm T})\big)_1$ holds the largest eigenvalue of symmetrized B, and so on. Then, for any $k\in\{1\ldots M\}$

$$\lambda\big(A+A^{\mathrm T}\big)_k + \lambda\big(B+B^{\mathrm T}\big)_M \;\leq\; \lambda\big((A+A^{\mathrm T})+(B+B^{\mathrm T})\big)_k \;\leq\; \lambda\big(A+A^{\mathrm T}\big)_k + \lambda\big(B+B^{\mathrm T}\big)_1 \tag{1593}$$

⋄

Weyl's theorem establishes: concavity of the smallest λ_M and convexity of the largest eigenvalue λ_1 of a symmetric matrix, via (496), and positive semidefiniteness of a sum of positive semidefinite matrices; for $A, B\in\mathbb{S}_+^M$

$$\lambda(A)_k + \lambda(B)_M \;\leq\; \lambda(A+B)_k \;\leq\; \lambda(A)_k + \lambda(B)_1 \tag{1594}$$

Because $\mathbb{S}_+^M$ is a convex cone (§2.9.0.0.1), then by (175)

$$A, B\succeq 0 \;\Rightarrow\; \zeta A+\xi B\succeq 0 \;\text{ for all }\; \zeta, \xi\geq 0 \tag{1595}$$

A.3.1.0.3 Corollary. *Eigenvalues of sum and difference.* [218, §4.3]
For $A\in\mathbb{S}^M$ and $B\in\mathbb{S}_+^M$, place the eigenvalues of each matrix into respective vectors $\lambda(A), \lambda(B)\in\mathbb{R}^M$ in nonincreasing order so $\lambda(A)_1$ holds the largest eigenvalue of A while $\lambda(B)_1$ holds the largest eigenvalue of B, and so on. Then, for any $k\in\{1\ldots M\}$

$$\lambda(A-B)_k \;\leq\; \lambda(A)_k \;\leq\; \lambda(A+B)_k \tag{1596}$$

⋄

When B is rank-one positive semidefinite, the eigenvalues interlace; *id est*, for $B=qq^{\mathrm T}$

$$\lambda(A)_{k-1} \;\leq\; \lambda(A-qq^{\mathrm T})_k \;\leq\; \lambda(A)_k \;\leq\; \lambda(A+qq^{\mathrm T})_k \;\leq\; \lambda(A)_{k+1} \tag{1597}$$

A.3.1.0.4 Theorem. *Positive (semi)definite principal submatrices.*[A.10]

- $A\in\mathbb{S}^M$ is positive semidefinite if and only if all M principal submatrices of dimension $M-1$ are positive semidefinite and $\det A$ is nonnegative.
- $A\in\mathbb{S}^M$ is positive definite if and only if any one principal submatrix of dimension $M-1$ is positive definite and $\det A$ is positive. ⋄

If any one principal submatrix of dimension $M-1$ is not positive definite, conversely, then A can neither be. Regardless of symmetry, if $A\in\mathbb{R}^{M\times M}$ is positive (semi)definite, then the determinant of each and every principal submatrix is (nonnegative) positive. [288, §1.3.1]

A.3.1.0.5 Corollary. *Positive (semi)definite symmetric products.* [218, p.399]

- If $A\in\mathbb{S}^M$ is positive definite and any particular dimensionally compatible matrix Z has no nullspace, then $Z^{\mathrm T}AZ$ is positive definite.
- If matrix $A\in\mathbb{S}^M$ is positive (semi)definite then, for any matrix Z of compatible dimension, $Z^{\mathrm T}AZ$ is positive semidefinite.
- $A\in\mathbb{S}^M$ is positive (semi)definite if and only if there exists a nonsingular Z such that $Z^{\mathrm T}AZ$ is positive (semi)definite.
- If $A\in\mathbb{S}^M$ is positive semidefinite and singular, then it remains possible that $Z^{\mathrm T}AZ$ becomes positive definite for some skinny $Z\in\mathbb{R}^{M\times N}$ $(N<M)$.[A.11] ⋄

We can deduce from these, given nonsingular matrix Z and any particular dimensionally compatible Y: matrix $A\in\mathbb{S}^M$ is positive semidefinite if and only if $\begin{bmatrix} Z^{\mathrm T} \\ Y^{\mathrm T} \end{bmatrix} A\,[\,Z \;\; Y\,]$ is positive semidefinite. In other words, from the Corollary it follows: for dimensionally compatible Z

[A.10] A recursive condition for positive (semi)definiteness, this theorem is a synthesis of facts from [218, §7.2] [348, §6.3] (*confer* [288, §1.3.1]).

[A.11] This means coefficients, of orthogonal projection of vectorized A on a subset of extreme directions from $\mathbb{S}_+^M$ determined by Z, can be positive (by the interpretation in §E.6.4.3).

- $A \succeq 0 \;\Leftrightarrow\; Z^{\mathrm T}AZ \succeq 0$ **and** $Z^{\mathrm T}$ has a left inverse

Products such as $Z^{\dagger}Z$ and $ZZ^{\dagger}$ are symmetric and positive semidefinite although, given $A \succeq 0$, $Z^{\dagger}AZ$ and $ZAZ^{\dagger}$ are neither necessarily symmetric or positive semidefinite.

A.3.1.0.6 Theorem. *Symmetric projector semidefinite.* [21, §III] [22, §6] [238, p.55] For symmetric idempotent matrices P and R

$$\begin{array}{c} P, R \succeq 0 \\ P \succeq R \;\Leftrightarrow\; \mathcal{R}(P) \supseteq \mathcal{R}(R) \;\Leftrightarrow\; \mathcal{N}(P) \subseteq \mathcal{N}(R) \end{array} \tag{1598}$$

Projector P is never positive definite [350, §6.5 prob.20] unless it is the Identity matrix. ⋄

A.3.1.0.7 Theorem. *Symmetric positive semidefinite.* [218, p.400]
Given real matrix Ψ with $\operatorname{rank}\Psi = 1$

$$\Psi \succeq 0 \;\Leftrightarrow\; \Psi = uu^{\mathrm T} \tag{1599}$$

where u is some real vector; *id est*, symmetry is necessary and sufficient for positive semidefiniteness of a rank-1 matrix. ⋄

Proof. Any rank-one matrix must have the form $\Psi = uv^{\mathrm T}$. (§B.1) Suppose Ψ is symmetric; *id est*, $v = u$. For all $y \in \mathbb{R}^M$, $y^{\mathrm T}u\,u^{\mathrm T}y \geq 0$. Conversely, suppose $uv^{\mathrm T}$ is positive semidefinite. We know that can hold if and only if $uv^{\mathrm T} + vu^{\mathrm T} \succeq 0 \;\Leftrightarrow\;$ for all normalized $y \in \mathbb{R}^M$, $2\,y^{\mathrm T}u\,v^{\mathrm T}y \geq 0$; but that is possible only if $v = u$. ♦

The same does not hold true for matrices of higher rank, as Example A.2.1.0.1 shows.

A.4 Schur complement

Consider *Schur-form* partitioned matrix G: Given $A^{\mathrm T} = A$ and $C^{\mathrm T} = C$, then [61]

$$\begin{array}{rl} G = \begin{bmatrix} A & B \\ B^{\mathrm T} & C \end{bmatrix} \succeq 0 & \\ \Leftrightarrow\; A \succeq 0\,, \;\; B^{\mathrm T}(I - AA^{\dagger}) = \mathbf{0}\,, \;\; C - B^{\mathrm T}A^{\dagger}B \succeq 0 \\ \Leftrightarrow\; C \succeq 0\,, \;\; B(I - CC^{\dagger}) = \mathbf{0}\,, \;\; A - B\,C^{\dagger}B^{\mathrm T} \succeq 0 \end{array} \tag{1600}$$

where $A^{\dagger}$ denotes the Moore-Penrose (pseudo)inverse (§E). In the first instance, $I - AA^{\dagger}$ is a symmetric projection matrix orthogonally projecting on $\mathcal{N}(A^{\mathrm T})$. (2013) It is apparently required

$$\mathcal{R}(B) \perp \mathcal{N}(A^{\mathrm T}) \tag{1601}$$

which precludes $A = \mathbf{0}$ when B is any nonzero matrix. Note that $A \succ 0 \Rightarrow A^{\dagger} = A^{-1}$; thereby, the projection matrix vanishes. Likewise, in the second instance, $I - CC^{\dagger}$ projects orthogonally on $\mathcal{N}(C^{\mathrm T})$. It is required

$$\mathcal{R}(B^{\mathrm T}) \perp \mathcal{N}(C^{\mathrm T}) \tag{1602}$$

which precludes $C=\mathbf{0}$ for B nonzero. Again, $C\succ 0 \;\Rightarrow\; C^{\dagger}=C^{-1}$. So we get, for A or C nonsingular,

$$\begin{array}{c} G=\left[\begin{array}{cc} A & B \\ B^{\mathrm{T}} & C \end{array}\right]\succeq 0 \\ \Leftrightarrow \\ A\succ 0\,,\quad C-B^{\mathrm{T}}A^{-1}B\succeq 0 \\ \textbf{or} \\ C\succ 0\,,\quad A-B\,C^{-1}B^{\mathrm{T}}\succeq 0 \end{array} \tag{1603}$$

When A is full-rank then, for all B of compatible dimension, $\mathcal{R}(B)$ is in $\mathcal{R}(A)$. Likewise, when C is full-rank, $\mathcal{R}(B^{\mathrm{T}})$ is in $\mathcal{R}(C)$. Thus the flavor, for A and C nonsingular,

$$\begin{array}{c} G=\left[\begin{array}{cc} A & B \\ B^{\mathrm{T}} & C \end{array}\right]\succ 0 \\ \Leftrightarrow\; A\succ 0\,,\quad C-B^{\mathrm{T}}A^{-1}B\succ 0 \\ \Leftrightarrow\; C\succ 0\,,\quad A-B\,C^{-1}B^{\mathrm{T}}\succ 0 \end{array} \tag{1604}$$

where $C-B^{\mathrm{T}}A^{-1}B$ is called the *Schur complement of* A *in* G, while the *Schur complement of* C *in* G is $A-B\,C^{-1}B^{\mathrm{T}}$. [167, §4.8]

Origin of the term *Schur complement* is from complementary *inertia*: [120, §2.4.4] Define

$$\operatorname{inertia}\Big(G\in\mathbb{S}^M\Big)\triangleq\{p\,,z\,,n\} \tag{1605}$$

where $p\,,z\,,n$ respectively represent number of positive, zero, and negative eigenvalues of G; *id est*,

$$M=p+z+n \tag{1606}$$

Then, when A is invertible,

$$\operatorname{inertia}(G)=\operatorname{inertia}(A)+\operatorname{inertia}(C-B^{\mathrm{T}}A^{-1}B) \tag{1607}$$

and when C is invertible,

$$\operatorname{inertia}(G)=\operatorname{inertia}(C)+\operatorname{inertia}(A-B\,C^{-1}B^{\mathrm{T}}) \tag{1608}$$

A.4.0.0.1 Example. *Equipartition inertia.* [56, §1.2 exer.17]
When $A=C=\mathbf{0}$, denoting nonincreasingly ordered singular values of matrix $B\in\mathbb{R}^{m\times m}$ by $\sigma(B)\in\mathbb{R}_+^m$, then we have eigenvalues

$$\lambda(G)\;=\;\lambda\left(\left[\begin{array}{cc} \mathbf{0} & B \\ B^{\mathrm{T}} & \mathbf{0} \end{array}\right]\right)\;=\;\left[\begin{array}{c} \sigma(B) \\ -\Xi\,\sigma(B) \end{array}\right] \tag{1609}$$

and

$$\operatorname{inertia}(G)=\operatorname{inertia}(B^{\mathrm{T}}B)+\operatorname{inertia}(-B^{\mathrm{T}}B) \tag{1610}$$

where Ξ is the order-reversing permutation matrix defined in (1828). □

A.4.0.0.2 Example. *Nonnegative polynomial.* [35, p.163]
Quadratic multivariate polynomial $x^{\mathrm T}A\,x + 2b^{\mathrm T}x + c$ is a convex function of vector x if and only if $A\succeq 0$, but sublevel set $\{x \mid x^{\mathrm T}A\,x + 2b^{\mathrm T}x + c \leq 0\}$ is convex if $A\succeq 0$ yet not *vice versa.* Schur-form positive semidefiniteness is sufficient for polynomial convexity but necessary and sufficient for nonnegativity:

$$\begin{bmatrix} A & b \\ b^{\mathrm T} & c \end{bmatrix} \succeq 0 \;\Leftrightarrow\; [\,x^{\mathrm T}\;\;1\,]\begin{bmatrix} A & b \\ b^{\mathrm T} & c \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix} \geq 0 \;\Leftrightarrow\; x^{\mathrm T}A\,x + 2b^{\mathrm T}x + c \geq 0 \tag{1611}$$

All is extensible to univariate polynomials; *e.g*, $x \triangleq [\,t^n \;\; t^{n-1} \;\; t^{n-2} \cdots t\,]^{\mathrm T}$. □

A.4.0.0.3 Example. *Schur-form fractional function trace minimization.*
From (1573),

$$\begin{array}{ccc}
\begin{bmatrix} A & B \\ B^{\mathrm T} & C \end{bmatrix} \succeq 0 & \Rightarrow & \operatorname{tr}(A + C) \geq 0 \\
\Updownarrow & & \\
\begin{bmatrix} A & \mathbf 0 \\ \mathbf 0^{\mathrm T} & C - B^{\mathrm T}A^{-1}B \end{bmatrix} \succeq 0 & \Rightarrow & \operatorname{tr}(C - B^{\mathrm T}A^{-1}B) \geq 0 \\
\Updownarrow & & \\
\begin{bmatrix} A - B\,C^{-1}B^{\mathrm T} & \mathbf 0 \\ \mathbf 0^{\mathrm T} & C \end{bmatrix} \succeq 0 & \Rightarrow & \operatorname{tr}(A - B\,C^{-1}B^{\mathrm T}) \geq 0
\end{array} \tag{1612}$$

Since $\operatorname{tr}(C - B^{\mathrm T}A^{-1}B) \geq 0 \;\Leftrightarrow\; \operatorname{tr} C \geq \operatorname{tr}(B^{\mathrm T}A^{-1}B) \geq 0$ for example, then minimization of $\operatorname{tr} C$ is necessary and sufficient for minimization of $\operatorname{tr}(C - B^{\mathrm T}A^{-1}B)$ when both are under constraint $\begin{bmatrix} A & B \\ B^{\mathrm T} & C \end{bmatrix} \succeq 0$. □

A.4.0.1 Schur-form nullspace basis

From (1600),

$$\begin{array}{ccc}
G = \begin{bmatrix} A & B \\ B^{\mathrm T} & C \end{bmatrix} \succeq 0 & & \\
\Updownarrow & & \\
\begin{bmatrix} A & \mathbf 0 \\ \mathbf 0^{\mathrm T} & C - B^{\mathrm T}A^{\dagger}B \end{bmatrix} \succeq 0 & \textbf{and} & B^{\mathrm T}(I - AA^{\dagger}) = \mathbf 0 \\
\Updownarrow & & \\
\begin{bmatrix} A - B\,C^{\dagger}B^{\mathrm T} & \mathbf 0 \\ \mathbf 0^{\mathrm T} & C \end{bmatrix} \succeq 0 & \textbf{and} & B(I - CC^{\dagger}) = \mathbf 0
\end{array} \tag{1613}$$

These facts plus Moore-Penrose condition (§E.0.1) provide a partial basis:

$$\operatorname{basis}\mathcal N\!\left(\begin{bmatrix} A & B \\ B^{\mathrm T} & C \end{bmatrix}\right) \supseteq \begin{bmatrix} I - AA^{\dagger} & \mathbf 0 \\ \mathbf 0^{\mathrm T} & I - CC^{\dagger} \end{bmatrix} \tag{1614}$$

A.4.0.1.1 Example. *Sparse Schur conditions.*
Setting matrix A to the Identity simplifies the Schur conditions. One consequence relates definiteness of three quantities:

$$\begin{bmatrix} I & B \\ B^{\mathrm{T}} & C \end{bmatrix} \succeq 0 \quad \Leftrightarrow \quad C - B^{\mathrm{T}}B \succeq 0 \quad \Leftrightarrow \quad \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & C - B^{\mathrm{T}}B \end{bmatrix} \succeq 0 \tag{1615}$$

□

A.4.0.1.2 Exercise. *Eigenvalues λ of sparse Schur-form.*
Prove: given $C - B^{\mathrm{T}}B = \mathbf{0}$, for $B \in \mathbb{R}^{m\times n}$ and $C \in \mathbb{S}^n$

$$\lambda\left(\begin{bmatrix} I & B \\ B^{\mathrm{T}} & C \end{bmatrix}\right)_i = \begin{cases} 1 + \lambda(C)_i\,, & 1 \le i \le n \\ 1\,, & n < i \le m \\ 0\,, & \text{otherwise} \end{cases} \tag{1616}$$

▼

A.4.0.1.3 Theorem. *Rank of partitioned matrices.* [432, §2.2 prob.7]
When symmetric matrix A is invertible and C is symmetric,

$$\begin{aligned} \operatorname{rank}\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} &= \operatorname{rank}\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & C - B^{\mathrm{T}}A^{-1}B \end{bmatrix} \\ &= \operatorname{rank} A \;+\; \operatorname{rank}(C - B^{\mathrm{T}}A^{-1}B) \end{aligned} \tag{1617}$$

equals rank of main diagonal block A plus rank of its Schur complement. Similarly, when symmetric matrix C is invertible and A is symmetric,

$$\begin{aligned} \operatorname{rank}\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} &= \operatorname{rank}\begin{bmatrix} A - B\,C^{-1}B^{\mathrm{T}} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & C \end{bmatrix} \\ &= \operatorname{rank}(A - B\,C^{-1}B^{\mathrm{T}}) + \operatorname{rank} C \end{aligned} \tag{1618}$$

⋄

Proof. The first assertion (1617) holds if and only if [218, §0.4.6c]

$$\exists \text{ nonsingular } X, Y \;\ni\; X\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} Y = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & C - B^{\mathrm{T}}A^{-1}B \end{bmatrix} \tag{1619}$$

Let [218, §7.7.6]

$$Y \;=\; X^{\mathrm{T}} = \begin{bmatrix} I & -A^{-1}B \\ \mathbf{0}^{\mathrm{T}} & I \end{bmatrix} \tag{1620}$$

◆

From Corollary A.3.1.0.3, eigenvalues are related by

$$0 \preceq \lambda(C - B^{\mathrm{T}}A^{-1}B) \preceq \lambda(C) \tag{1621}$$

$$0 \preceq \lambda(A - B\,C^{-1}B^{\mathrm{T}}) \preceq \lambda(A) \tag{1622}$$

which means

$$\operatorname{rank}(C - B^{\mathrm{T}}A^{-1}B) \le \operatorname{rank} C \tag{1623}$$

$$\operatorname{rank}(A - B\,C^{-1}B^{\mathrm{T}}) \le \operatorname{rank} A \tag{1624}$$

Therefore

$$\operatorname{rank}\begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} \le \operatorname{rank} A + \operatorname{rank} C \tag{1625}$$

A.4.0.1.4 Lemma. *Rank of Schur-form block.* [147] [145]
Matrix $B \in \mathbb{R}^{m\times n}$ has $\operatorname{rank} B \le \rho$ if and only if there exist matrices $A \in \mathbb{S}^m$ and $C \in \mathbb{S}^n$ such that

$$\operatorname{rank}\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & C \end{bmatrix} \le 2\rho \qquad \textbf{and} \qquad G = \begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} \succeq 0 \tag{1626}$$

⋄

Schur-form positive semidefiniteness alone implies $\operatorname{rank} A \ge \operatorname{rank} B$ and $\operatorname{rank} C \ge \operatorname{rank} B$. But, even in absence of semidefiniteness, we must always have $\operatorname{rank} G \ge \operatorname{rank} A\,,\ \operatorname{rank} B\,,\ \operatorname{rank} C$ by fundamental linear algebra.

A.4.1 Determinant

$$G = \begin{bmatrix} A & B \\ B^{\mathrm{T}} & C \end{bmatrix} \tag{1627}$$

We consider again a matrix G partitioned like (1600), but not necessarily positive (semi)definite, where A and C are symmetric.

- When A is invertible,

$$\det G = \det A \ \det(C - B^{\mathrm{T}}A^{-1}B) \tag{1628}$$

When C is invertible,

$$\det G = \det C \ \det(A - B\,C^{-1}B^{\mathrm{T}}) \tag{1629}$$

- When B is full-rank and skinny, $C = \mathbf{0}$, and $A \succeq 0$, then [63, §10.1.1]

$$\det G \ne 0 \ \Leftrightarrow \ A + BB^{\mathrm{T}} \succ 0 \tag{1630}$$

When B is a (column) vector, then for all $C \in \mathbb{R}$ and all A of dimension compatible with G

$$\det G = \det(A)C - B^{\mathrm{T}}A_{\mathrm{cof}}^{\mathrm{T}}B \tag{1631}$$

while for $C \ne 0$

$$\det G = C \det(A - \frac{1}{C}BB^{\mathrm{T}}) \tag{1632}$$

where A_{cof} is the matrix of cofactors [348, §4] corresponding to A.

- When B is full-rank and fat, $A=\mathbf{0}$, and $C\succeq 0$, then

$$\det G\neq 0 \;\Leftrightarrow\; C+B^{\mathrm T}B\succ 0 \tag{1633}$$

When B is a row-vector, then for $A\neq 0$ and all C of dimension compatible with G

$$\det G \;=\; A\,\det(C-\frac{1}{A}B^{\mathrm T}B) \tag{1634}$$

while for all $A\in\mathbb{R}$

$$\det G \;=\; \det(C)A-B\,C_{\mathrm{cof}}^{\mathrm T}B^{\mathrm T} \tag{1635}$$

where C_{cof} is the matrix of cofactors corresponding to C.

A.5 Eigenvalue decomposition

All square matrices have associated eigenvalues λ and eigenvectors; if not square, $Ax=\lambda_i x$ becomes impossible dimensionally. Eigenvectors must be nonzero. Prefix *eigen* is from the German; in this context meaning, something akin to "characteristic". [345, p.14]

When a square matrix $X\in\mathbb{R}^{m\times m}$ is *diagonalizable,* [348, §5.6] then

$$X \;=\; S\Lambda S^{-1} = [\,s_1\cdots s_m\,]\;\Lambda\begin{bmatrix} w_1^{\mathrm T}\\ \vdots\\ w_m^{\mathrm T}\end{bmatrix} = \sum_{i=1}^{m}\lambda_i\,s_i w_i^{\mathrm T} \tag{1636}$$

where $\{s_i\in\mathcal{N}(X-\lambda_i I)\subseteq\mathbb{C}^m\}$ are l.i. (right-)eigenvectors constituting the columns of $S\in\mathbb{C}^{m\times m}$ defined by

$$XS=S\Lambda \qquad \textbf{rather} \qquad Xs_i\triangleq\lambda_i s_i\;, \quad i=1\ldots m \tag{1637}$$

$\{w_i\in\mathcal{N}(X^{\mathrm T}-\lambda_i I)\subseteq\mathbb{C}^m\}$ are linearly independent *left-eigenvectors* of X (eigenvectors of $X^{\mathrm T}$) constituting the rows of S^{-1} defined by [218]

$$S^{-1}X=\Lambda S^{-1} \qquad \textbf{rather} \qquad w_i^{\mathrm T}X\triangleq\lambda_i w_i^{\mathrm T}\;, \quad i=1\ldots m \tag{1638}$$

and where $\{\lambda_i\in\mathbb{C}\}$ are eigenvalues (1550)

$$\delta(\lambda(X))=\Lambda\in\mathbb{C}^{m\times m} \tag{1639}$$

corresponding to both left and right eigenvectors; *id est*, $\lambda(X)=\lambda(X^{\mathrm T})$.

There is no connection between diagonalizability and invertibility of X. [348, §5.2] Diagonalizability is guaranteed by a full set of linearly independent eigenvectors, whereas invertibility is guaranteed by all nonzero eigenvalues.

$$\begin{array}{c}\text{distinct eigenvalues} \;\Rightarrow\; \text{l.i. eigenvectors} \;\Leftrightarrow\; \text{diagonalizable}\\ \text{not diagonalizable} \;\Rightarrow\; \text{repeated eigenvalue}\end{array} \tag{1640}$$

$\begin{bmatrix} 1 & 0 & -1\\ -1 & 1 & 0\\ 3 & -1 & -2\end{bmatrix}$ is not diagonalizable, for example, having three 0-eigenvalues which are hard to compute with accuracy better than `1E-6`. (Yates, D'Errico)

A.5.0.0.1 Theorem. *Real eigenvector.*
Eigenvectors of a real matrix corresponding to real eigenvalues must be real. ⋄

Proof. $Ax = \lambda x$. Given $\lambda = \lambda^*$, $x^{\mathrm{H}}Ax = \lambda x^{\mathrm{H}}x = \lambda\|x\|^2 = x^{\mathrm{T}}A\,x^* \Rightarrow x = x^*$, where $x^{\mathrm{H}} = x^{*\mathrm{T}}$. The converse is equally simple. ♦

A.5.0.1 Uniqueness

From the *fundamental theorem of algebra,* [235] which guarantees existence of zeros for a given polynomial, it follows: eigenvalues, including their multiplicity, for a given square matrix are unique; meaning, there is no other set of eigenvalues for that matrix. (Conversely, many different matrices may share the same unique set of eigenvalues; *e.g*, for any X, $\lambda(X) = \lambda(X^{\mathrm{T}})$.)

Uniqueness of eigenvectors, in contrast, disallows multiplicity of the same direction:

A.5.0.1.1 Definition. *Unique eigenvectors.*
When eigenvectors are *unique,* we mean: unique to within a real nonzero scaling, and their directions are distinct. △

If S is a matrix of eigenvectors of X as in (1636), for example, then $-S$ is certainly another matrix of eigenvectors decomposing X with the same eigenvalues. Although directions are distinct, eigenvectors $-S$ are equivalent to eigenvectors S by Definition A.5.0.1.1.

For any square matrix, the eigenvector corresponding to a distinct eigenvalue is unique; [345, p.220]

$$\text{distinct eigenvalues} \Rightarrow \text{eigenvectors unique} \tag{1641}$$

Eigenvectors corresponding to a repeated eigenvalue are not unique for a diagonalizable matrix;

$$\text{repeated eigenvalue} \Rightarrow \text{eigenvectors not unique} \tag{1642}$$

> Proof follows from the observation: any linear combination of distinct eigenvectors of diagonalizable X, corresponding to a particular eigenvalue, produces another eigenvector. For eigenvalue λ whose multiplicity[A.12] $\dim\mathcal{N}(X-\lambda I)$ exceeds 1, in other words, any choice of independent vectors from $\mathcal{N}(X-\lambda I)$ (of the same multiplicity) constitutes eigenvectors corresponding to λ. ♦

Caveat diagonalizability insures linear independence which implies existence of distinct eigenvectors. We may conclude, for diagonalizable matrices,

$$\text{distinct eigenvalues} \Leftrightarrow \text{eigenvectors unique} \tag{1643}$$

[A.12] A matrix is diagonalizable iff *algebraic multiplicity* (number of occurrences of same eigenvalue) equals *geometric multiplicity* $\dim\mathcal{N}(X-\lambda I) = m - \operatorname{rank}(X-\lambda I)$ [345, p.15] (number of *Jordan blocks* w.r.t λ or number of corresponding l.i. eigenvectors).

A.5.0.2 Invertible matrix

When diagonalizable matrix $X \in \mathbb{R}^{m\times m}$ is *nonsingular* (no zero eigenvalues), then it has an inverse obtained simply by inverting eigenvalues in (1636):

$$X^{-1} = S\Lambda^{-1}S^{-1} \tag{1644}$$

A.5.0.3 eigenmatrix

The (right-)eigenvectors $\{s_i\}$ (1636) are naturally orthogonal $w_i^{\mathrm T}s_j = 0$ to left-eigenvectors $\{w_i\}$ except, for $i = 1 \ldots m$, $w_i^{\mathrm T}s_i = 1$; called a *biorthogonality condition* [384, §2.2.4] [218] because neither set of left or right eigenvectors is necessarily an orthogonal set. Consequently, each dyad from a diagonalization is an independent (§B.1.1) nonorthogonal projector because

$$s_i w_i^{\mathrm T} s_i w_i^{\mathrm T} = s_i w_i^{\mathrm T} \tag{1645}$$

(whereas the dyads of singular value decomposition are not inherently projectors (*confer* (1652))).

Dyads of eigenvalue decomposition can be termed *eigenmatrices* because

$$X\, s_i w_i^{\mathrm T} = \lambda_i\, s_i w_i^{\mathrm T} \tag{1646}$$

Sum of the eigenmatrices is the Identity;

$$\sum_{i=1}^{m} s_i w_i^{\mathrm T} = I \tag{1647}$$

A.5.1 Symmetric matrix diagonalization

The set of *normal matrices* is, precisely, that set of all real matrices having a complete orthonormal set of eigenvectors; [432, §8.1] [350, prob.10.2.31] *id est*, any matrix X for which $XX^{\mathrm T} = X^{\mathrm T}X$; [174, §7.1.3] [345, p.3] *e.g*, symmetric, orthogonal, and circulant matrices [184]. All normal matrices are diagonalizable.

A symmetric matrix is a special normal matrix whose eigenvalues Λ must be real[A.13] and whose eigenvectors S can be chosen to make a real orthonormal set; [350, §6.4] [348, p.315] *id est*, for $X \in \mathbb{S}^m$

$$X \;=\; S\Lambda S^{\mathrm T} \;=\; [\, s_1 \cdots s_m \,]\; \Lambda \begin{bmatrix} s_1^{\mathrm T} \\ \vdots \\ s_m^{\mathrm T} \end{bmatrix} \;=\; \sum_{i=1}^{m} \lambda_i\, s_i s_i^{\mathrm T} \tag{1648}$$

where $\delta^2(\Lambda) = \Lambda \in \mathbb{S}^m$ (§A.1) and $S^{-1} = S^{\mathrm T} \in \mathbb{R}^{m\times m}$ (orthogonal matrix, §B.5.2) because of symmetry: $S\Lambda S^{-1} = S^{-\mathrm T}\Lambda S^{\mathrm T}$. By 0 *eigenvalues theorem* A.7.3.0.1,

$$\begin{array}{rcl} \mathcal{R}\{s_i \,|\, \lambda_i \neq 0\} &=& \mathcal{R}(A) \;=\; \mathcal{R}(A^{\mathrm T}) \\ \mathcal{R}\{s_i \,|\, \lambda_i = 0\} &=& \mathcal{N}(A^{\mathrm T}) \;=\; \mathcal{N}(A) \end{array} \tag{1649}$$

[A.13] **Proof.** Suppose λ_i is an eigenvalue corresponding to eigenvector s_i of real $A = A^{\mathrm T}$. Then $s_i^{\mathrm H}As_i = s_i^{\mathrm T}As_i^*$ (by transposition) $\Rightarrow$ $s_i^{*\mathrm T}\lambda_i s_i = s_i^{\mathrm T}\lambda_i^* s_i^*$ because $(As_i)^* = (\lambda_i s_i)^*$ by assumption. So we have $\lambda_i\|s_i\|^2 = \lambda_i^*\|s_i\|^2$. There is no converse. ♦

A.5.1.1 Diagonal matrix diagonalization

Because arrangement of eigenvectors and their corresponding eigenvalues is arbitrary, we almost always arrange eigenvalues in nonincreasing order as is the convention for singular value decomposition. Then to diagonalize a symmetric matrix that is already a diagonal matrix, orthogonal matrix S becomes a permutation matrix.

A.5.1.2 Invertible symmetric matrix

When symmetric matrix $X\in\mathbb{S}^m$ is nonsingular (invertible), then its inverse (obtained by inverting eigenvalues in (1648)) is also symmetric:

$$X^{-1} = S\Lambda^{-1}S^{\mathrm{T}}\in\mathbb{S}^m \tag{1650}$$

A.5.1.3 Positive semidefinite matrix square root

When $X\in\mathbb{S}_+^m$, its unique positive semidefinite matrix square root is defined

$$\sqrt{X} \triangleq S\sqrt{\Lambda}\,S^{\mathrm{T}} \in \mathbb{S}_+^m \tag{1651}$$

where the square root of nonnegative diagonal matrix $\sqrt{\Lambda}$ is taken entrywise and positive. Then $X=\sqrt{X}\sqrt{X}$.

A.6 Singular value decomposition, SVD

A.6.1 Compact SVD

[174, §2.5.4] For any $A\in\mathbb{R}^{m\times n}$

$$\begin{gathered} A \;=\; U\Sigma Q^{\mathrm{T}} \;=\; [\,u_1 \cdots u_\eta\,]\;\Sigma\begin{bmatrix} q_1^{\mathrm{T}} \\ \vdots \\ q_\eta^{\mathrm{T}} \end{bmatrix} = \sum_{i=1}^{\eta}\sigma_i\,u_i\,q_i^{\mathrm{T}} \\ U\in\mathbb{R}^{m\times\eta}, \quad \Sigma\in\mathbb{R}_+^{\eta\times\eta}, \quad Q\in\mathbb{R}^{n\times\eta} \\ U^{\mathrm{T}}U = I\,, \qquad\qquad Q^{\mathrm{T}}Q = I \end{gathered} \tag{1652}$$

where U and Q are always skinny-or-square real, each having orthonormal columns, and where

$$\eta \triangleq \min\{m\,,\,n\} \tag{1653}$$

Square matrix Σ is diagonal (§A.1.1)

$$\delta^2(\Sigma) = \Sigma\in\mathbb{R}_+^{\eta\times\eta} \tag{1654}$$

holding the singular values $\{\sigma_i \in \mathbb{R}\}$ of A which are always arranged in nonincreasing order by convention and are related to eigenvalues λ by[A.14]

$$\sigma(A)_i = \sigma(A^{\mathrm{T}})_i = \begin{cases} \sqrt{\lambda(A^{\mathrm{T}}A)_i} \;=\; \sqrt{\lambda(AA^{\mathrm{T}})_i} \;=\; \lambda\Big(\sqrt{A^{\mathrm{T}}A}\Big)_i = \lambda\Big(\sqrt{AA^{\mathrm{T}}}\Big)_i > 0\,, & 1 \leq i \leq \rho \\ 0\,, & \rho < i \leq \eta \end{cases} \tag{1655}$$

of which the last $\eta - \rho$ are 0 ,[A.15] where

$$\rho \triangleq \operatorname{rank} A \;=\; \operatorname{rank} \Sigma \tag{1656}$$

A point sometimes lost: Any real matrix may be decomposed in terms of its real singular values $\sigma(A) \in \mathbb{R}^\eta$ and real matrices U and Q as in (1652), where [174, §2.5.3]

$$\begin{array}{l} \mathcal{R}\{u_i \mid \sigma_i \neq 0\} = \mathcal{R}(A) \\ \mathcal{R}\{u_i \mid \sigma_i = 0\} \subseteq \mathcal{N}(A^{\mathrm{T}}) \\ \mathcal{R}\{q_i \mid \sigma_i \neq 0\} = \mathcal{R}(A^{\mathrm{T}}) \\ \mathcal{R}\{q_i \mid \sigma_i = 0\} \subseteq \mathcal{N}(A) \end{array} \tag{1657}$$

A.6.2 Subcompact SVD

Some authors allow only nonzero singular values. In that case the compact decomposition can be made smaller; it can be redimensioned in terms of rank ρ because, for any $A \in \mathbb{R}^{m\times n}$

$$\rho \;=\; \operatorname{rank} A \;=\; \operatorname{rank} \Sigma \;=\; \max\{i \in \{1 \ldots \eta\} \mid \sigma_i \neq 0\} \;\leq\; \eta \tag{1658}$$

- There are η singular values. For any flavor SVD, rank is equivalent to the number of nonzero singular values on the main diagonal of Σ .

Now

$$\begin{array}{c} A \;=\; U\Sigma Q^{\mathrm{T}} = [\, u_1 \cdots u_\rho \,]\; \Sigma \begin{bmatrix} q_1^{\mathrm{T}} \\ \vdots \\ q_\rho^{\mathrm{T}} \end{bmatrix} = \sum_{i=1}^{\rho} \sigma_i\, u_i q_i^{\mathrm{T}} \\ U \in \mathbb{R}^{m\times\rho}, \quad \Sigma \in \mathbb{R}_+^{\rho\times\rho}, \quad Q \in \mathbb{R}^{n\times\rho} \\ U^{\mathrm{T}}U = I\,, \qquad\qquad Q^{\mathrm{T}}Q = I \end{array} \tag{1659}$$

where the main diagonal of diagonal matrix Σ has no 0 entries, and

$$\begin{array}{l} \mathcal{R}\{u_i\} = \mathcal{R}(A) \\ \mathcal{R}\{q_i\} = \mathcal{R}(A^{\mathrm{T}}) \end{array} \tag{1660}$$

[A.14] When matrix A is normal, $\sigma(A) = |\lambda(A)|$. [432, §8.1]

[A.15] For $\eta = n$, $\sigma(A) = \sqrt{\lambda(A^{\mathrm{T}}A)} = \lambda\Big(\sqrt{A^{\mathrm{T}}A}\Big)$.

For $\eta = m$, $\sigma(A) = \sqrt{\lambda(AA^{\mathrm{T}})} = \lambda\Big(\sqrt{AA^{\mathrm{T}}}\Big)$.

A.6.3 Full SVD

Another common and useful expression of the SVD makes U and Q square; making the decomposition larger than compact SVD. Completing the nullspace bases in U and Q from (1657) provides what is called the *full singular value decomposition* of $A\in\mathbb{R}^{m\times n}$ [348, App.A]. Orthonormal real matrices U and Q become orthogonal matrices (§B.5):

$$\begin{aligned}\mathcal{R}\{u_i \mid \sigma_i\neq 0\} &= \mathcal{R}(A)\\ \mathcal{R}\{u_i \mid \sigma_i= 0\} &= \mathcal{N}(A^{\mathrm{T}})\\ \mathcal{R}\{q_i \mid \sigma_i\neq 0\} &= \mathcal{R}(A^{\mathrm{T}})\\ \mathcal{R}\{q_i \mid \sigma_i= 0\} &= \mathcal{N}(A)\end{aligned} \tag{1661}$$

For any matrix A having rank ρ $(=\operatorname{rank}\Sigma)$

$$\begin{aligned}A &= U\Sigma Q^{\mathrm{T}} = [\,u_1 \cdots u_m\,]\;\Sigma\begin{bmatrix} q_1^{\mathrm{T}}\\ \vdots\\ q_n^{\mathrm{T}}\end{bmatrix} = \sum_{i=1}^{\eta}\sigma_i\, u_i q_i^{\mathrm{T}}\\ &= \big[\,m\times\rho\ \ \text{basis}\,\mathcal{R}(A) \qquad m\times m-\rho\ \ \text{basis}\,\mathcal{N}(A^{\mathrm{T}})\,\big]\begin{bmatrix}\sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & \end{bmatrix}\begin{bmatrix}\big(n\times\rho\ \ \text{basis}\,\mathcal{R}(A^{\mathrm{T}})\big)^{\mathrm{T}}\\ \big(n\times n-\rho\ \ \text{basis}\,\mathcal{N}(A)\big)^{\mathrm{T}}\end{bmatrix}\\ &\qquad U\in\mathbb{R}^{m\times m},\quad \Sigma\in\mathbb{R}_+^{m\times n},\quad Q\in\mathbb{R}^{n\times n}\\ &\qquad U^{\mathrm{T}}=U^{-1}\,, \qquad\qquad Q^{\mathrm{T}}=Q^{-1}\end{aligned} \tag{1662}$$

where upper limit of summation η is defined in (1653). Matrix Σ is no longer necessarily square, now padded with respect to (1654) by $m-\eta$ zero rows or $n-\eta$ zero columns; the nonincreasingly ordered (possibly 0) singular values appear along its main diagonal as for compact SVD (1655).

*An important geometrical interpretation of SVD is given in Figure **169** for $m=n=2$: The image of the unit sphere under any $m\times n$ matrix multiplication is an ellipse. Considering the three factors of the SVD separately, note that Q^{T} is a pure rotation of the circle. Figure **169** shows how the axes q_1 and q_2 are first rotated by Q^{T} to coincide with the coordinate axes. Second, the circle is stretched by Σ in the directions of the coordinate axes to form an ellipse. The third step rotates the ellipse by U into its final position. Note how q_1 and q_2 are rotated to end up as u_1 and u_2, the principal axes of the final ellipse. A direct calculation shows that $Aq_j=\sigma_j\,u_j$. Thus q_j is first rotated to coincide with the j^{th} coordinate axis, stretched by a factor σ_j, and then rotated to point in the direction of u_j. All of this is beautifully illustrated for 2×2 matrices by the* MATLAB *code* `eigshow.m` (see [351]).

A direct consequence of the geometric interpretation is that the largest singular value σ_1 measures the "magnitude" of A (its 2-norm):

$$\|A\|_2 = \sup_{\|x\|_2=1}\|Ax\|_2 = \sigma_1 \tag{1663}$$

This means that $\|A\|_2$ is the length of the longest principal semiaxis of the ellipse.

$$A = U\Sigma Q^{\mathrm{T}} = [\, u_1 \cdots u_m \,]\ \Sigma \begin{bmatrix} q_1^{\mathrm{T}} \\ \vdots \\ q_n^{\mathrm{T}} \end{bmatrix} = \sum_{i=1}^{\eta} \sigma_i\, u_i q_i^{\mathrm{T}}$$

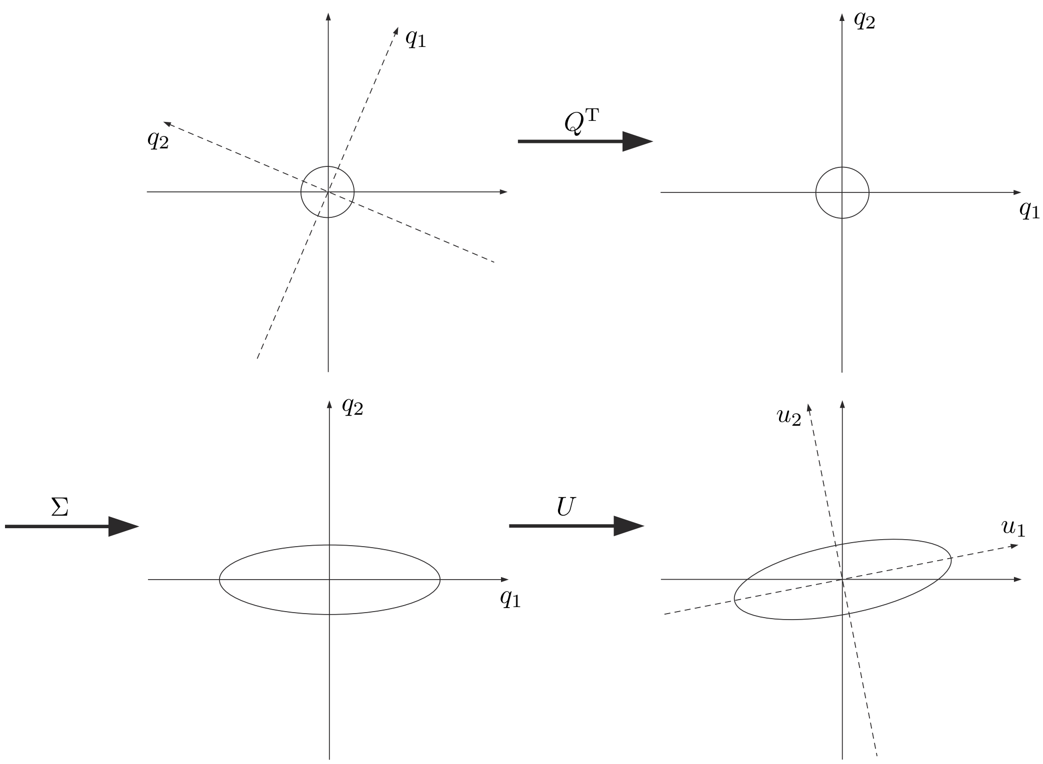

Figure 169: Full SVD geometrical interpretation [287]: Image of circle $\{x \in \mathbb{R}^{\mathbf{2}} \mid \|x\|_2 = 1\}$, under matrix multiplication Ax, is generally an ellipse. For the example illustrated, $U \triangleq [\, u_1 \;\; u_2 \,] \in \mathbb{R}^{\mathbf{2}\times\mathbf{2}}$, $Q \triangleq [\, q_1 \;\; q_2 \,] \in \mathbb{R}^{\mathbf{2}\times\mathbf{2}}$.

Expressions for U , Q , *and* Σ *follow readily from* (1662),

$$AA^{\mathrm T}U = U\Sigma\Sigma^{\mathrm T} \quad\text{and}\quad A^{\mathrm T}AQ = Q\Sigma^{\mathrm T}\Sigma \tag{1664}$$

demonstrating that the columns of U *are the eigenvectors of* $AA^{\mathrm T}$ *and the columns of* Q *are the eigenvectors of* $A^{\mathrm T}A$. −Muller, Magaia, & Herbst [287]

A.6.4 Pseudoinverse by SVD

Matrix pseudoinverse (§E) is nearly synonymous with singular value decomposition because of the elegant expression, given $A = U\Sigma Q^{\mathrm T}\in\mathbb{R}^{m\times n}$

$$A^{\dagger} = Q\Sigma^{\dagger\mathrm T}U^{\mathrm T}\in\mathbb{R}^{n\times m} \tag{1665}$$

that applies to all three flavors of SVD, where $\Sigma^{\dagger}$ simply inverts nonzero entries of matrix Σ .

Given symmetric matrix $A\in\mathbb{S}^n$ and its diagonalization $A = S\Lambda S^{\mathrm T}$ (§A.5.1), its pseudoinverse simply inverts all nonzero eigenvalues:

$$A^{\dagger} = S\Lambda^{\dagger}S^{\mathrm T}\in\mathbb{S}^n \tag{1666}$$

A.6.5 SVD of symmetric matrices

From (1655) and (1651) for $A = A^{\mathrm T}$

$$\sigma(A)_i = \begin{cases} \sqrt{\lambda(A^2)_i} \;=\; \lambda\Big(\sqrt{A^2}\Big)_i \;=\; |\lambda(A)_i| \;>\; 0\,, & 1\le i\le\rho \\ 0\,, & \rho < i\le\eta \end{cases} \tag{1667}$$

A.6.5.0.1 Definition. *Step function.* (*confer* §4.3.2.0.1)
Define the signum-like quasilinear function $\psi : \mathbb{R}^n\to\mathbb{R}^n$ that takes value 1 corresponding to a 0-valued entry in its argument:

$$\psi(a) \triangleq \left[\lim_{x_i\to a_i}\frac{x_i}{|x_i|} \;=\; \begin{cases} 1\,, & a_i\ge 0 \\ -1\,, & a_i<0 \end{cases}\;,\quad i{=}1\ldots n\right]\in\mathbb{R}^n \tag{1668}$$

△

Eigenvalue signs of a symmetric matrix having diagonalization $A = S\Lambda S^{\mathrm T}$ (1648) can be absorbed either into real U or real Q from the full SVD; [369, p.34] (*confer* §C.4.2.1)

$$A \;=\; S\Lambda S^{\mathrm T} \;=\; S\delta(\psi(\delta(\Lambda)))\,|\Lambda|S^{\mathrm T} \triangleq U\,\Sigma Q^{\mathrm T}\in\mathbb{S}^n \tag{1669}$$

or

$$A \;=\; S\Lambda S^{\mathrm T} \;=\; S|\Lambda|\,\delta(\psi(\delta(\Lambda)))S^{\mathrm T} \triangleq U\Sigma\,Q^{\mathrm T}\in\mathbb{S}^n \tag{1670}$$

where matrix of singular values $\Sigma = |\Lambda|$ denotes entrywise absolute value of diagonal eigenvalue matrix Λ .

A.7 Zeros

A.7.1 norm zero

For any given norm, by definition,

$$\|x\|_\ell = 0 \quad \Leftrightarrow \quad x = \mathbf{0} \tag{1671}$$

Consequently, a generally nonconvex constraint in x like $\|Ax - b\| = \kappa$ becomes convex when $\kappa = 0$.

A.7.2 0 entry

If a positive semidefinite matrix $A = [A_{ij}] \in \mathbb{R}^{n\times n}$ has a 0 entry A_{ii} on its main diagonal, then $A_{ij} + A_{ji} = 0 \;\; \forall j$. [288, §1.3.1]

Any symmetric positive semidefinite matrix having a 0 entry on its main diagonal must be $\mathbf{0}$ along the entire row and column to which that 0 entry belongs. [174, §4.2.8] [218, §7.1 prob.2] From which it follows: for $A \in \mathbb{S}^n_+$

$$\delta(A) = \mathbf{0} \;\Leftrightarrow\; A = \mathbf{0} \tag{1672}$$

$$\operatorname{tr}(A) = 0 \;\Leftrightarrow\; A = \mathbf{0} \tag{1673}$$

A.7.3 *0 eigenvalues theorem*

This theorem is simple, powerful, and widely applicable:

A.7.3.0.1 Theorem. *Number of* 0 *eigenvalues.*
For any matrix $A \in \mathbb{R}^{m\times n}$

$$\operatorname{rank}(A) + \dim \mathcal{N}(A) = n \tag{1674}$$

by conservation of dimension. [218, §0.4.4]

For any square matrix $A \in \mathbb{R}^{m\times m}$, number of 0 eigenvalues is at least equal to $\dim \mathcal{N}(A)$

$$\dim \mathcal{N}(A) \;\leq\; \text{number of } 0 \text{ eigenvalues} \;\leq\; m \tag{1675}$$

while all eigenvectors corresponding to those 0 eigenvalues belong to $\mathcal{N}(A)$. [348, §5.1][A.16]

[A.16] We take as given the well-known fact that the number of 0 eigenvalues cannot be less than dimension of the nullspace. We offer an example of the converse:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$\dim \mathcal{N}(A) = 2$, $\lambda(A) = [\,0\;0\;0\;1\,]^{\mathrm T}$; three eigenvectors in the nullspace but only two are independent. The right-hand side of (1675) is tight for nonzero matrices; *e.g*, (§B.1) dyad $uv^{\mathrm T} \in \mathbb{R}^{m\times m}$ has m 0-eigenvalues when $u \in v^{\perp}$.

For diagonalizable matrix A (§A.5), the number of 0 eigenvalues is precisely $\dim\mathcal{N}(A)$ while the corresponding eigenvectors span $\mathcal{N}(A)$. Real and imaginary parts of the eigenvectors remaining span $\mathcal{R}(A)$.

(Transpose.)

Likewise, for any matrix $A\in\mathbb{R}^{m\times n}$

$$\operatorname{rank}(A^{\mathrm{T}}) + \dim\mathcal{N}(A^{\mathrm{T}}) = m \tag{1676}$$

For any square $A\in\mathbb{R}^{m\times m}$, number of 0 eigenvalues is at least equal to $\dim\mathcal{N}(A^{\mathrm{T}})=\dim\mathcal{N}(A)$ while all left-eigenvectors (eigenvectors of A^{T}) corresponding to those 0 eigenvalues belong to $\mathcal{N}(A^{\mathrm{T}})$.

For diagonalizable A, number of 0 eigenvalues is precisely $\dim\mathcal{N}(A^{\mathrm{T}})$ while the corresponding left-eigenvectors span $\mathcal{N}(A^{\mathrm{T}})$. Real and imaginary parts of the left-eigenvectors remaining span $\mathcal{R}(A^{\mathrm{T}})$. ⋄

Proof. First we show, for a diagonalizable matrix, the number of 0 eigenvalues is precisely the dimension of its nullspace while the eigenvectors corresponding to those 0 eigenvalues span the nullspace:

Any diagonalizable matrix $A\in\mathbb{R}^{m\times m}$ must possess a complete set of linearly independent eigenvectors. If A is full-rank (invertible), then all $m=\operatorname{rank}(A)$ eigenvalues are nonzero. [348, §5.1]

Suppose $\operatorname{rank}(A) < m$. Then $\dim\mathcal{N}(A) = m-\operatorname{rank}(A)$. Thus there is a set of $m-\operatorname{rank}(A)$ linearly independent vectors spanning $\mathcal{N}(A)$. Each of those can be an eigenvector associated with a 0 eigenvalue because A is diagonalizable ⇔ ∃ m linearly independent eigenvectors. [348, §5.2] Eigenvectors of a real matrix corresponding to 0 eigenvalues must be real.[A.17] Thus A has at least $m-\operatorname{rank}(A)$ eigenvalues equal to 0.

Now suppose A has more than $m-\operatorname{rank}(A)$ eigenvalues equal to 0. Then there are more than $m-\operatorname{rank}(A)$ linearly independent eigenvectors associated with 0 eigenvalues, and each of those eigenvectors must be in $\mathcal{N}(A)$. Thus there are more than $m-\operatorname{rank}(A)$ linearly independent vectors in $\mathcal{N}(A)$; a contradiction.

Diagonalizable A therefore has $\operatorname{rank}(A)$ nonzero eigenvalues and exactly $m-\operatorname{rank}(A)$ eigenvalues equal to 0 whose corresponding eigenvectors span $\mathcal{N}(A)$.

By similar argument, the left-eigenvectors corresponding to 0 eigenvalues span $\mathcal{N}(A^{\mathrm{T}})$.

Next we show when A is diagonalizable, the real and imaginary parts of its eigenvectors (corresponding to nonzero eigenvalues) span $\mathcal{R}(A)$:

The (right-)eigenvectors of a diagonalizable matrix $A\in\mathbb{R}^{m\times m}$ are linearly independent if and only if the left-eigenvectors are. So, matrix A has a representation in terms of its right- and left-eigenvectors; from the diagonalization (1636), assuming 0 eigenvalues are ordered last,

$$A = \sum_{i=1}^{m}\lambda_i\, s_i w_i^{\mathrm{T}} = \sum_{\substack{i=1\\ \lambda_i\neq 0}}^{k\,\leq\, m} \lambda_i\, s_i w_i^{\mathrm{T}} \tag{1677}$$

From the *linearly independent dyads theorem* (§B.1.1.0.2), the dyads $\{s_i w_i^{\mathrm{T}}\}$ must be independent because each set of eigenvectors are; hence $\operatorname{rank} A = k$, the number of nonzero

[A.17] **Proof.** Let * denote complex conjugation. Suppose $A=A^*$ and $As_i=\mathbf{0}$. Then $s_i=s_i^* \Rightarrow As_i=As_i^* \Rightarrow As_i^*=\mathbf{0}$. Conversely, $As_i^*=\mathbf{0} \Rightarrow As_i=As_i^* \Rightarrow s_i=s_i^*$. ♦

eigenvalues. Complex eigenvectors and eigenvalues are common for real matrices, and must come in complex conjugate pairs for the summation to remain real. Assume that conjugate pairs of eigenvalues appear in sequence. Given any particular conjugate pair from (1677), we get the partial summation

$$\begin{aligned} \lambda_i\, s_i w_i^{\mathrm T} + \lambda_i^*\, s_i^* w_i^{*\mathrm T} &= 2\,\mathrm{re}(\lambda_i\, s_i w_i^{\mathrm T}) \\ &= 2\big(\mathrm{re}\, s_i\ \mathrm{re}(\lambda_i\, w_i^{\mathrm T}) - \mathrm{im}\, s_i\ \mathrm{im}(\lambda_i\, w_i^{\mathrm T})\big) \end{aligned} \tag{1678}$$

where[A.18] $\lambda_i^* \triangleq \lambda_{i+1}$, $s_i^* \triangleq s_{i+1}$, and $w_i^* \triangleq w_{i+1}$. Then (1677) is equivalently written

$$A = 2 \sum_{\substack{i \\ \lambda\in\mathbb{C} \\ \lambda_i\neq 0}} \mathrm{re}\, s_{2i}\ \mathrm{re}(\lambda_{2i}\, w_{2i}^{\mathrm T}) - \mathrm{im}\, s_{2i}\ \mathrm{im}(\lambda_{2i}\, w_{2i}^{\mathrm T}) \ + \sum_{\substack{j \\ \lambda\in\mathbb{R} \\ \lambda_j\neq 0}} \lambda_j\, s_j w_j^{\mathrm T} \tag{1679}$$

The summation (1679) shows: A is a linear combination of real and imaginary parts of its (right-)eigenvectors corresponding to nonzero eigenvalues. The k vectors $\{\mathrm{re}\, s_i\in\mathbb{R}^m,\ \mathrm{im}\, s_i\in\mathbb{R}^m \mid \lambda_i\neq 0,\ i\in\{1\ldots m\}\}$ must therefore span the range of diagonalizable matrix A.

The argument is similar regarding span of the left-eigenvectors. ♦

A.7.4 0 trace and matrix product

For $X, A\in\mathbb{R}_+^{M\times N}$ (39)

$$\mathrm{tr}(X^{\mathrm T}A) = 0 \ \Leftrightarrow\ X\circ A = A\circ X = \mathbf{0} \tag{1680}$$

For $X, A\in\mathbb{S}_+^M$ [35, §2.6.1 exer.2.8] [378, §3.1]

$$\mathrm{tr}(XA) = 0 \ \Leftrightarrow\ XA = AX = \mathbf{0} \tag{1681}$$

Proof. ($\Leftarrow$) Suppose $XA = AX = \mathbf{0}$. Then $\mathrm{tr}(XA)=0$ is obvious.
($\Rightarrow$) Suppose $\mathrm{tr}(XA)=0$. $\mathrm{tr}(XA)=\mathrm{tr}(\sqrt{A}\,X\sqrt{A})$ whose argument is positive semidefinite by Corollary A.3.1.0.5. Trace of any square matrix is equivalent to the sum of its eigenvalues. Eigenvalues of a positive semidefinite matrix can total 0 if and only if each and every nonnegative eigenvalue is 0. The only positive semidefinite matrix, having all 0 eigenvalues, resides at the origin; (*confer* (1705)) *id est*,

$$\sqrt{A}\,X\sqrt{A} = \left(\sqrt{X}\sqrt{A}\right)^{\mathrm T}\sqrt{X}\sqrt{A} = \mathbf{0} \tag{1682}$$

implying $\sqrt{X}\sqrt{A}=\mathbf{0}$ which in turn implies $\sqrt{X}(\sqrt{X}\sqrt{A})\sqrt{A} = XA = \mathbf{0}$. Arguing similarly yields $AX=\mathbf{0}$. ♦

Diagonalizable matrices A and X are *simultaneously diagonalizable* if and only if they are commutative under multiplication; [218, §1.3.12] *id est*, iff they share a complete set of eigenvectors.

[A.18] Complex conjugate of w is denoted w^*. Conjugate transpose is denoted $w^{\mathrm H} = w^{*\mathrm T}$.

A.7.4.0.1 Example. *An equivalence in nonisomorphic spaces.*
Identity (1681) leads to an unusual equivalence relating convex geometry to traditional linear algebra: The convex sets, given $A \succeq 0$

$$\{X \mid \langle X\,,\,A\rangle = 0\} \cap \{X \succeq 0\} \;\equiv\; \{X \mid \mathcal{N}(X) \supseteq \mathcal{R}(A)\} \cap \{X \succeq 0\} \tag{1683}$$

(one expressed in terms of a hyperplane, the other in terms of nullspace and range) are equivalent only when symmetric matrix A is positive semidefinite.

We might apply this equivalence to the geometric center subspace, for example,

$$\begin{aligned} \mathbb{S}_c^M &= \{Y \in \mathbb{S}^M \mid Y\mathbf{1} = \mathbf{0}\} \\ &= \{Y \in \mathbb{S}^M \mid \mathcal{N}(Y) \supseteq \mathbf{1}\} \;=\; \{Y \in \mathbb{S}^M \mid \mathcal{R}(Y) \subseteq \mathcal{N}(\mathbf{1}^{\mathrm T})\} \end{aligned} \tag{2113}$$

from which we derive (*confer* (1082))

$$\mathbb{S}_c^M \cap \mathbb{S}_+^M \;\equiv\; \{X \succeq 0 \mid \langle X\,,\,\mathbf{1}\mathbf{1}^{\mathrm T}\rangle = 0\} \tag{1684}$$

□

A.7.5 Zero definite

The domain over which an arbitrary real matrix A is zero definite can exceed its left and right nullspaces. For any positive semidefinite matrix $A \in \mathbb{R}^{M\times M}$ (for $A + A^{\mathrm T} \succeq 0$)

$$\{x \mid x^{\mathrm T}\!Ax = 0\} \;=\; \mathcal{N}(A + A^{\mathrm T}) \tag{1685}$$

because $\exists\, R \ni A + A^{\mathrm T} = R^{\mathrm T}R$, $\|Rx\| = 0 \Leftrightarrow Rx = \mathbf{0}$, and $\mathcal{N}(A + A^{\mathrm T}) = \mathcal{N}(R)$. Then given any particular vector $x_{\mathrm p}$, $x_{\mathrm p}^{\mathrm T}Ax_{\mathrm p} = 0 \;\Leftrightarrow\; x_{\mathrm p} \in \mathcal{N}(A + A^{\mathrm T})$. For any positive definite matrix A (for $A + A^{\mathrm T} \succ 0$)

$$\{x \mid x^{\mathrm T}\!Ax = 0\} \;=\; \mathbf{0} \tag{1686}$$

Further, [432, §3.2 prob.5]

$$\{x \mid x^{\mathrm T}\!Ax = 0\} = \mathbb{R}^M \;\Leftrightarrow\; A^{\mathrm T} = -A \tag{1687}$$

while

$$\{x \mid x^{\mathrm H}\!Ax = 0\} = \mathbb{C}^M \;\Leftrightarrow\; A = \mathbf{0} \tag{1688}$$

The positive semidefinite matrix

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \tag{1689}$$

for example, has no nullspace. Yet

$$\{x \mid x^{\mathrm T}\!Ax = 0\} \;=\; \{x \mid \mathbf{1}^{\mathrm T}x = 0\} \subset \mathbb{R}^{\mathbf{2}} \tag{1690}$$

which is the nullspace of the symmetrized matrix. Symmetric matrices are not spared from the excess; *videlicet*,

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \tag{1691}$$

has eigenvalues $\{-1, 3\}$, no nullspace, but is zero definite on[A.19]

$$\mathcal{X} \triangleq \{x \in \mathbb{R}^2 \mid x_2 = (-2 \pm \sqrt{3}\,)x_1\} \tag{1692}$$

A.7.5.0.1 Proposition. (Sturm/Zhang) *Dyad-decompositions.* [354, §5.2]
Let positive semidefinite matrix $X \in \mathbb{S}_+^M$ have rank ρ. Then, given symmetric matrix $A \in \mathbb{S}^M$, $\langle A\,, X\rangle = 0$ if and only if there exists a dyad-decomposition

$$X = \sum_{j=1}^{\rho} x_j x_j^{\mathrm{T}} \tag{1693}$$

satisfying

$$\langle A\,,\, x_j x_j^{\mathrm{T}}\rangle = 0 \quad \text{for each and every } j \in \{1 \ldots \rho\} \tag{1694}$$

$\diamond$

The dyad-decomposition of X proposed is generally not that obtained from a standard diagonalization by eigenvalue decomposition, unless $\rho = 1$ or the given matrix A is simultaneously diagonalizable (§A.7.4) with X. That means, elemental dyads in decomposition (1693) constitute a generally nonorthogonal set. Sturm & Zhang give a simple procedure for constructing the dyad-decomposition [407] where matrix A may be regarded as a parameter.

A.7.5.0.2 Example. *Dyad.*
The dyad $uv^{\mathrm{T}} \in \mathbb{R}^{M\times M}$ (§B.1) is zero definite on all x for which either $x^{\mathrm{T}}u = 0$ or $x^{\mathrm{T}}v = 0$;

$$\{x \mid x^{\mathrm{T}}uv^{\mathrm{T}}x = 0\} \;=\; \{x \mid x^{\mathrm{T}}u = 0\} \cup \{x \mid v^{\mathrm{T}}x = 0\} \tag{1695}$$

id est, on $u^{\perp} \cup v^{\perp}$. Symmetrizing the dyad does not change the outcome:

$$\{x \mid x^{\mathrm{T}}(uv^{\mathrm{T}} + vu^{\mathrm{T}})x/2 = 0\} \;=\; \{x \mid x^{\mathrm{T}}u = 0\} \cup \{x \mid v^{\mathrm{T}}x = 0\} \tag{1696}$$

$\square$

[A.19] These two lines represent the limit in the union of two generally distinct hyperbolae; *id est*, for matrix B and set $\mathcal{X}$ as defined

$$\lim_{\varepsilon \to 0^+} \{x \in \mathbb{R}^2 \mid x^{\mathrm{T}}Bx = \varepsilon\} \;=\; \mathcal{X}$$