9. Positive definite mappings

First, we'll introduce the tensor product, which we'll use later in this section.

Given two vector spaces U, V over K, their tensor product, denoted $U \otimes V$, is a related vector space of dimension $(\dim U)(\dim V)$.

Def 1: If $\{u_1, \dots, u_n\}$ is a basis of U and $\{v_1, \dots, v_m\}$ a basis of V, then $\{u_i \otimes v_j : 1 \le i \le n,\ 1 \le j \le m\}$ is a basis of $U \otimes V$.

Analogy: $U \cong \text{Span}\{1, x, \dots, x^{n-1}\}$, $V \cong \text{Span}\{1, y, \dots, y^{m-1}\}$;

$U \otimes V \cong \text{Span}\{x^i y^j : 0 \le i < n,\ 0 \le j < m\}$.

Compare to $U \times V$, which has dimension $\dim U + \dim V$, with basis $\{(x_i, 0)\} \cup \{(0, y_j)\}$.

There's a better, basis-free way to construct $U \otimes V$.

Def 2: $U \otimes V = \{\sum u_i \otimes v_i : u_i \in U, v_i \in V\} / N$

where $N = \mathrm{Span}\{ (u_1+u_2)\otimes v - u_1 \otimes v - u_2 \otimes v,$
$u \otimes (v_1+v_2) - u\otimes v_1 - u \otimes v_2 \}$

Basically, we are forcing the distributive law, i.e,

$(u_1+u_2)\otimes v = u_1 \otimes v + u_2 \otimes v$, etc.

Let $\mathrm{Hom}(X,Y)$ be the set of linear maps from X to Y.

Theorem 9.1: There is a natural isomorphism $U\otimes V \longrightarrow \mathrm{Hom}(U', V)$.

Proof (sketch). Define the map as follows:

$$U \otimes V \longrightarrow \mathrm{Hom}(U', V)$$

$$u \otimes v \longmapsto \{ \ell \mapsto (\ell, u)\, v \}$$ and extend linearily

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \otimes \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \longmapsto \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} [u_1 \cdots u_n] = \begin{bmatrix} v_1u_1 & v_1u_2 & \cdots & v_1u_n \\ \vdots & & \ddots & \vdots \\ v_mu_1 & v_mv_2 & \cdots & v_mu_n \end{bmatrix}_{m\times n}$$

Remark: We could similarly define an isomorphism $U\otimes V \to \mathrm{Hom}(V', U)$.

Also, $e_i \otimes e_j$ under this isomorphism is the matrix E_{ij} (i.e, the ij-entry is 1, all others 0)

If U, V are Euclidean spaces (so $U' = U$), there is a natural way to endow $U \otimes V$ with a Euclidean structure.

For $M, L \in \mathcal{L}(U,V)$, define $(M, L) = tr(L^* M) = \sum_{i,j} \bar{l}_{ji}\, m_{ji}$.

Note that $\|M\|^2 = (M, M) = \sum_{i,j} |m_{ji}|^2$

Clearly, $\{E_{ij} : 1 \le i \le n,\ 1 \le j \le m\}$ is an orthonormal basis.

Ex: $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $L = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$, all real.

$$(M, L) = tr\, L^* M = tr \begin{bmatrix} a' & c' \\ b' & d' \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$= tr \begin{bmatrix} aa' + cc' & a'b + c'd \\ ab' + cd' & bb' + dd' \end{bmatrix}$$

$$= aa' + bb' + cc' + dd'.$$

Another way to view tensor products.

Let X be an n-dimensional real vector space.

Note that $\mathbb{C}$ is a 2-dim'l $\mathbb{R}$-vector space (basis $\{1, i\}$).

Suppose $A : X \to X$ is a linear map with min'l poly x^2+1.

Then i and $-i$ are eigenvalues of A, but $i \notin \mathbb{R}$.

So if v is an eigenvector for $\lambda = i$, $v \notin X$.

However, v should live in some "extension" of X.

In this bigger vector space, we want to have vectors like zv, $z \in \mathbb{C}$, $v \in X$.

What we really want is $\mathbb{C} \otimes X$.

This has basis $\{x_1, \dots, x_n, ix_1, \dots, ix_n\}$, where $x_1, \dots, x_n$ is a basis of X [and here, $ix_j \leftrightarrow i \otimes x_j$]

Note that we need certain associativity and distributivity, like $(3i)v = (i3)v = i(3v)$

i.e., $3i \otimes v = i3 \otimes v = i \otimes 3v$.

But this comes for free with the construction!

Similarly, compare this to polynomials & matrices:

$(3x^i)y^j = x^i(3y^j)$ and $(3u)v^T = u(3v^T)$

$3x^i \otimes y^j = x^i \otimes 3y^j$ $\qquad$ $3u \otimes v = u \otimes 3v$.

<u>Def:</u> A self-adjoint map $H: X \to X$ is <u>positive</u> (or positive definite) if $(x, Hx) > 0$ for all $x \neq 0$. It is <u>nonnegative</u> (or positive semidefinite) if $(x, Hx) \geq 0$.

We denote these as e.g. $H > 0$ and $H \geq 0$, resp.

<u>Theorem 9.2</u>: Let X be a Euclidean space.

(i) The identity map I is positive.

(ii) If $M, N > 0$ then $M + N > 0$ and $aM > 0$ for $a > 0$.

(iii) If $H > 0$ and Q inv't, then $Q^* H Q > 0$.

(iv) $H > 0$ iff all eigenvalues are positive.

(v) If $H > 0$ then H is invertible.

(vi) Every positive map has a unique positive square root.

(vii) The set of positive maps is an open subset of the space of self-adjoint maps.

(viii) The boundary points of the set of positive maps are the nonnegative maps that are not positive.

$\boxed{6}$

Proof: (i) $(x, Ix) = (x, x) > 0$ if $x \neq 0$. ✓

(ii) $(x, (M+N)x) = (x, Mx) + (x, Nx) > 0$ if $x \neq 0$. ✓

and $(x, aMx) = a(x, Mx) > 0$ ✓

(iii) $(x, Q^* H Q x) = (Qx, HQx)$.

$= (y, Hy) > 0$ where $y = Qx$ ✓

(iv) $(\Rightarrow)$ $(x, Hx) = (x, \lambda x) = \lambda(x, x) = 0 \Rightarrow \lambda > 0$

$(\Leftarrow)$ $\dfrac{(x, Hx)}{(x, x)} \geq \lambda_{min} > 0 \Rightarrow (x, Hx) \geq \lambda_{min} \|x\| > 0$. ✓

(v) If H is singular, then $Hx = 0$ for some $x \neq 0 \Rightarrow \lambda = 0$. ↯

(vi) Write $\sqrt{H}\, x = \sum_{i=1}^{n} a_i \sqrt{\lambda_i}\, x_i$ (x_i's e-vectors). ✓

(vii) Fix $H > 0$, let N be any self-adjoint map such that $\|N - H\| < \lambda_{min}$.

Claim: $N > 0$.

Put $M = N - H$. Since $\|M\| < \lambda_{min}$, $\|Mx\| < \lambda_{min} \|x\|$ for all $x \neq 0$.

Cauchy-Schwarz $\Rightarrow |(x, Mx)| \leq \|x\| \cdot \|Mx\| < \lambda_{min} \|x\|^2$.

Together we get for $x \neq 0$:

$$(x, Nx) = (x, (H+M)x) = \underbrace{(x, Hx)}_{\geq \lambda_{min}\|x\|^2} + \underbrace{(x, Mx)}_{> -\lambda_{min}\|x\|^2} > \lambda_{min}\|x\|^2 - \lambda_{min}\|x\|^2 = 0.$$

Thus $N > 0$. ✓

(viii) By definition, if K is on the boundary of positive maps, then $\exists$ sequence $H_n \to K$ of positive maps i.e., $\|H_n - K\| \to 0$.

By Cauchy-Schwarz, $\lim_{n\to\infty} (x, H_n x) = (x, Kx)$, so $(x, Kx) \geq 0$.

Since $K \not> 0$, then $K \geq 0$. ✓ $\square$

Put a <u>partial order</u> onto the set of <u>self-adjoint maps</u>:

Say $M < N$ iff $N - M > 0$

$M \leq N$ iff $N - M \geq 0$.

We get the following properties (almost) for free:

Additive: $M_1 < N_1$ and $M_2 < N_2 \Rightarrow M_1 + M_2 < N_1 + N_2$

Transitive: $L < M < N \Rightarrow L < N$

Multiplicative: $M < N$, Q inv't $\Rightarrow Q^* M Q < Q^* N Q$.

<u>Theorem 9.3</u>: Suppose $0 < M < N$. Then $M^{-1} > N^{-1} > 0$.

<u>Proof</u>: First, suppose $N = I$.

Then $M < I \Rightarrow I - M > 0 \Rightarrow$ Eigenvalues of $I - M$ are positive.

Say $(I - M)x = \lambda x$ for $x \neq 0$.

Then $Mx = x - \lambda x = (1 - \lambda)x$. Since $1 - \lambda > 0$, $0 < \lambda < 1$.

In summary: λ eigenvalue of $I-M \Rightarrow 0<\lambda<1$

$\Rightarrow 1-\lambda$ eigenvalue of M, and $0<1-\lambda<1$

$\Rightarrow \frac{1}{1-\lambda}$ eigenvalue of M^{-1}, and $\frac{1}{1-\lambda}>1$

$\Rightarrow$ Eigenvalues of $M^{-1}-I>0$

$\Rightarrow M^{-1}-I>0 \Rightarrow M^{-1}>I.$ ✓

Now, consider arbitrary $N>M>0$.

Factor $N=R^2$, $R>0$ and invertible.

Put $\underline{Q=R^{-1}}$: (so $Q^*=R^{-1}$)

$$0<M<N \Rightarrow 0<R^{-1}MR^{-1}<R^{-1}NR^{-1}=I$$

Take inverses: $0>RM^{-1}R>I$

$$\Rightarrow 0>R^{-1}(RM^{-1}R)R^{-1}>R^{-1}IR^{-1}=R^{-2}=N^{-1}$$

$$\Rightarrow 0>M^{-1}>N^{-1}. \qquad \square$$

<u>Caveat</u>: The product of self-adjoint maps is in general, not self-adjoint.

<u>Example</u>: Let $A=\begin{pmatrix}1&2\\2&5\end{pmatrix}$, $B=\begin{pmatrix}1&-2\\-2&5\end{pmatrix}$, $x=\begin{pmatrix}1\\0\end{pmatrix}$.

$Ax=\begin{pmatrix}1\\2\end{pmatrix}$, $Bx=\begin{pmatrix}1\\-2\end{pmatrix}$, $(x, ABx)=(Ax, Bx)=-3$.

Def: If A, B are self-adjoint, define their symmetrized product as $S = AB + BA$.

Note that $(x, Sx) = (x, ABx) + (x, BAx) = (Ax, Bx) + (Bx, Ax)$.

In the real case, $(x, Sx) = 2(Ax, Bx)$.

In the example above, $AB + BA = \begin{pmatrix} -6 & 0 \\ 0 & 42 \end{pmatrix}$.

That is, it's false that $A > 0$, $B > 0 \Rightarrow AB + BA > 0$.

But a similar statement is true:

Theorem 9.4: Let A, B be self-adjoint. If $A > 0$ and $AB + BA > 0$, then $B > 0$.

Proof: Define $B(t) = B + tA$. (Note: We must show $B(0) > 0$.)

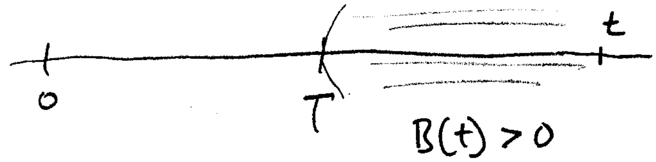
Claim 1: The symmetrized product of A & $B(t)$ is positive for $t \geq 0$.

$$S(t) = A\,B(t) + B(t)\,A = A(B + tA) + (B + tA)A = \underbrace{AB + BA}_{= S > 0} + \underbrace{2tA^2}_{> 0} = S + 2tA^2 > 0$$

Claim 2: $\exists\ T$ s.t. if $t > T$, then $B(t) > 0$.

$$(x, B(t)x) = (x, (B + tA)x) = (x, Bx) + t(x, Ax) \qquad [\text{assume } \|x\| = 1]$$

Recall:
- $(x, Ax) \geq \lambda_{min} \|x\|^2 = \lambda_{min}$
- $|(x, Bx)| \leq \|x\| \cdot \|Bx\| \leq \|B\| \cdot \|x\|^2 = \|B\|$



Together, if $\|x\| = 1$, then

$$(x, B(t)x) = t(x, Ax) + (x, Bx)$$
$$\geq t\lambda_{min} - \|B\| > 0 \text{ if } t > \|B\|/\lambda_{min} := T$$

Claim 3: $B = B(0) > 0$.

If not, then for some t_0, $0 \leq t_0 \leq \|B\|/\lambda_{min}$ such that

$B(t_0)$ is on the boundary of the set of positive maps,

i.e., $B(t_0) \geq 0$ but $B(t_0) \not> 0$.

Such a $B(t_0)$ has a $\lambda = 0$, so $B(t_0)y = 0$ for some $y \neq 0$.

However, $(y, S(t_0)y) = (Ay, B(t_0)y) + (B(t_0)y, Ay) = 0$. $\lightning$

Thus $B > 0$. $\square$

Corollary 9.5: If $0 < M < N$, then $0 < \sqrt{M} < \sqrt{N}$.

Proof: Put $A(t) = M + t(N-M)$.

For $0 \leq t \leq 1$, $A(t) = (1-t)M + tN > 0$, $\dot{A}(t) = N - M > 0$.

Thus, we can define $R(t) = \sqrt{A(t)}$ for $0 \leq t \leq 1$.

Since $A = R^2$, $\dot{A} = \dot{R}R + R\dot{R}$ (symmetrized product of R & $\dot{R}$)

[See Lax Ch. 9 for the "product rule" of linear map derivatives]

We know $\dot{A} = N - M > 0 \Rightarrow \dot{R} > 0$ on $[0,1]$.

Claim: $R(t)$ is an increasing function on $[0,1]$.

For any $x \neq 0$, $\frac{d}{dt}(x, Rx) = (x, \dot{R}x) > 0$.

By calculus (see Lax Ch. 9), $(x, R(t)x)$ is increasing.

$$\Rightarrow (x, R(s)x) < (x, R(t)x) \quad \text{for } s < t$$

$$\Rightarrow R(t) - R(s) > 0 \Rightarrow R(t) > R(s).$$

In particular $R(0) < R(1)$

and $R(0) = \sqrt{A(0)} = \sqrt{M}$, $R(1) = \sqrt{A(1)} = \sqrt{N} \Rightarrow \sqrt{M} < \sqrt{N}$. $\square$

Def: A real-valued function $f(s)$, $s > 0$ is a monotone matrix function (mmf) if for all self-adjoint mappings,

$$0 < M < N \Rightarrow f(M) < f(N).$$

Recall that $f(H) = \sum_{i=1}^{n} f(\lambda_i) P_i$ where $H = \sum_{i=1}^{n} \lambda_i P_i$.

Examples:

(i) $f(s) = -\frac{1}{s}$ is a mmf (immediate from Theorem 9.3;

$$0 < M < N \Rightarrow M^{-1} > N^{-1} > 0 \Rightarrow -M^{-1} < -N^{-1} < 0.)$$

(ii) $f(s) = s^{1/2}$ is a mmf (by Corollary 10.5).

(iii) $f(s) = s^2$ is <u>not</u> a mmf.

Take any $A, B > 0$ with $S = AB + BA \ngeq 0$.

<u>Claim</u>: For small t, if $M = A$, $N = A + tB$, then

$0 < M < N$ but $M^2 \nless N^2$.

<u>why</u>: $N^2 = A^2 + \underbrace{t(AB + BA)}_{\text{not negligible}} + \underbrace{t^2 B}_{\text{negligible}}$ for $t \approx 0$

So $N^2 = M^2 + [\text{something non-positive}] \ngeq M^2$

(iv) $f(s) = s^{(-2^k)}$ and $f(s) = \log s$ are mmf. (HW).

Additionally, positive multiples, sums, and limits of mmf's are mmf's.

For example, $-\sum \frac{m_j}{s + t_j}$ $m_j > 0$, $t_j > 0$ is an mmf, and

so is $f(s) = as + b - \int_0^\infty \frac{dm(t)}{s + t}$ $a > 0$, $b \in \mathbb{R}$, $(*)$

and $m(t)$ non-negative measure for which integral converges.

In fact, <u>every</u> mmf has the form of $(**)$

(Theorem of C. Loewner — very non-trivial!).

Surprisingly, functions of the form (*) are easy to characterize:

__Theorem__ (Herglotz, Riesz): Every function f which is analytic on the upper half-plane with $\mathrm{Im}(f) > 0$ there, and $\mathrm{Im}(f) = 0$ on the positive real axis, has the form (*).

Conversely, __every__ function of the form (*) can be extended to be analytic on the upper half-plane, with $\mathrm{Im}(f) > 0$ there.

__Proof__. See Lax's book "Functional analysis."

How to construct (all) positive matrices:

__Def__. Let $f_1, \dots, f_m$ be a sequence of vectors in a Euclidean space X. Define the $m \times m$ matrix G, where $G_{ij} = (f_i, f_j)$.
This is called the __Gram matrix__ of $f_1, \dots, f_m$.

__Theorem 9.6__:

(i) Every Gram matrix is nonnegative.

(ii) The Gram matrix of a set of linearly independent vectors is positive

(iii) Every positive matrix is a Gram matrix!

Proof: (i), (ii): $(x, Gx) = \sum_{i,j} x_i \overline{G_{ij}}\, \bar{x}_j = \sum_{i,j} (f_i, f_j)\, x_i \bar{x}_j$

$$= \Big(\sum_{i=1}^{m} x_i f_i , \sum_{j=1}^{m} x_j f_j \Big) = \Big\| \sum x_i f_i \Big\|^2$$

(iii) Let $H = (h_{ij})$ be positive, and define the <u>nonstandard inner product</u> by $(x, y)_H := (x, Hy)$.

Note that the Gram matrix of $e_1, \dots, e_m$ has ij-entry

$$(e_i, e_j)_H = (e_i, He_j) = h_{ij}. \qquad \square$$

<u>Examples:</u>

(i) Let $X = \{ f : [0,1] \to \mathbb{R} \}$, $(f, g) := \int_0^1 f(t)\, g(t)\, dt$.

If $f_1 = 1$, $f_2 = t, \dots, f_i = t^{i-1}$, then

$G = (G_{ij})$ where $G_{ij} = \dfrac{1}{i+j-1}$.

(ii) Define $(f, g) = \int_0^{2\pi} f(\theta)\, \overline{g(\theta)}\, w(\theta)\, d\theta$, $w : \mathbb{R} \to \mathbb{R}^+$.

If $f_j = e^{ij\theta}$, $j = -n, \dots, n$, then the $(2n+1) \times (2n+1)$ corresponding Gram matrix is $G_{kj} = c_{k-j}$, where

$$c_p = \int w(\theta)\, e^{-ip\theta}\, d\theta.$$

Theorem 9.7 (Schur): Let $A = (A_{ij})$ and $B = (B_{ij})$ be positive matrices. Then $M = (M_{ij}) := (A_{ij} B_{ij})$ is positive.

Proof: Since A & B are Gram matrices (thm 9.6) write $A_{ij} = (u_i, u_j)$, $B_{ij} = (v_i, v_j)$ where $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are linearly independent.

Define $g_i \in U \otimes V$ as $g_i := u_i \otimes v_i$.

Note that $(g_i, g_j) = (u_i, u_j)(v_i, v_j) = A_{ij} B_{ij}$ (Exercise)

Thus M is a Gram matrix, and so $M > 0$ by Thm 9.6. $\square$

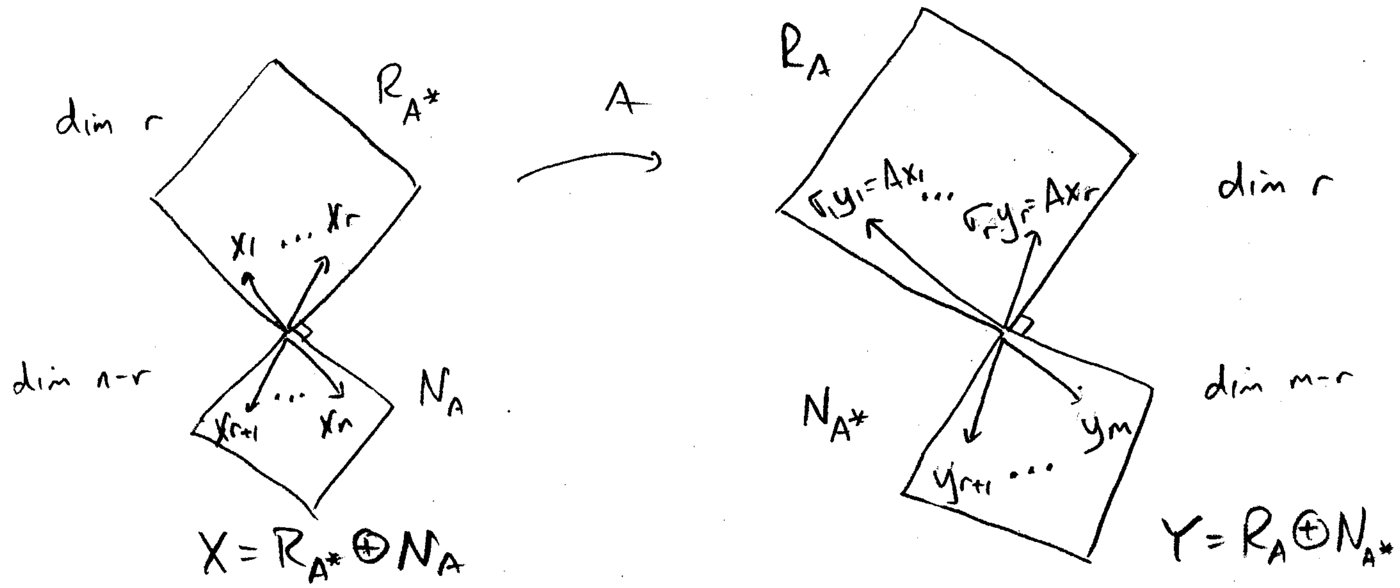
Singular value decomposition (SVD)

Big idea: If $A : X \to Y$ is linear, write $\boxed{A = U \Sigma V^*}$, where U, V are orthogonal, Σ diagonal

Special Case: $A = Q^* D Q$ if A is self-adjoint

No good: $A = P^{-1} D P$ (e-vectors might not be orthog.)

"Cartoon" of this:



Goal: Find orthonormal bases of R_{A^*} and R_A so

$$A\begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix} = \begin{bmatrix} y_1 & y_2 & \cdots & y_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$$

Once we have this, we can extend these bases to bases of all of X & Y,

$$A\Big[\underbrace{x_1 \cdots x_r}_{\text{basis of } R_{A^*}} \ \underbrace{x_{r+1} \cdots x_n}_{\text{basis of } N_A}\Big] = \Big[\underbrace{y_1 \cdots y_r}_{\text{basis of } R_A} \ \underbrace{y_{r+1} \cdots y_m}_{\text{basis of } N_{A^*}}\Big]\begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Question: How to find U, V, Σ ?

(It's not even clear why these should exist!)

Recall that A^*A and AA^* are self-adjoint (real e-values) and nonnegative (all e-values ≥ 0).

$$A^*A = (V\Sigma^* U^*)(U\Sigma V^*) = V\begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \end{bmatrix}V^*.$$

$$AA^* = (U\Sigma V^*)(V\Sigma^* U^*) = U\begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \ddots \end{bmatrix}U^*.$$

So: V = matrix of eigenvectors of A^*A

U = matrix of eigenvectors of AA^*

Σ = matrix of square roots of eigenvalues of A^*A (or AA^*).

Pseudoinverses

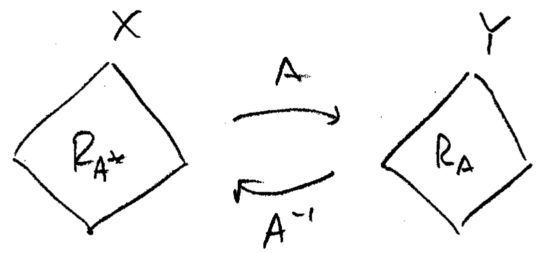
Let $A: X \to Y$ be linear.

Recall: $X = R_{A^*} \oplus N_A \xrightarrow{A} R_A \oplus N_{A^*} = Y$

and the restriction $R_{A^*} \xrightarrow{A} R_A$ is a bijection.

Big idea: Even if A is noninvertible, it may have a left, right, or pseudoinverse.

Case 1: A has a 2-sided inverse:

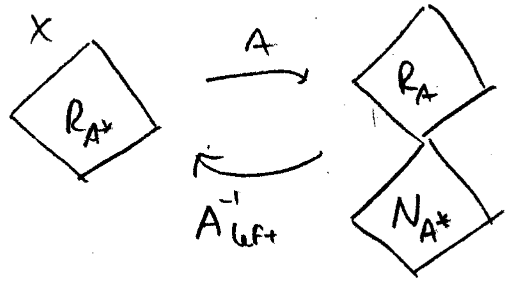


$AA^{-1} = I = A^{-1}A$

full rank $r = n = m$

$N_A = N_{A^*} = \{0\}$

$Ax = b$ has 1 solution



Case 2: A has a left-inverse.

$N_A = \{0\}$ "full column rank"

$Ax = b$ has 0 or 1 solution.

A^*A is invertible:

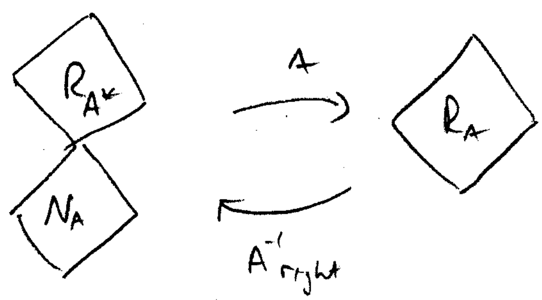
$$\underbrace{(A^*A)^{-1}A^*}_{A^{-1}_{left}} A = I_{n\times n}$$

Reverse order:

$$AA^{-1}_{left} = A\left[(A^*A)^{-1}A^*\right]$$

projection onto R_A.

Case 3: A has a right-inverse.



$N_{A^*} = \{0\}$ "full row rank"

$Ax = b$ has ∞ solutions

AA^* is invertible

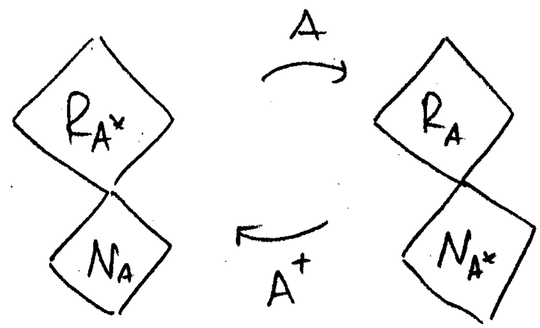
$$A\underbrace{A^*(AA^*)^{-1}}_{A^{-1}_{right}} = I_{m\times m}$$

Reverse order:

$$A^{-1}_{right}A = \left[A^*(AA^*)^{-1}\right]A$$

projection onto R_{A^*}

Case 4 The "general case" (any A!)



Want: Matrix A^+ "the pseudoinverse"

such that

$$AA^+ = \begin{bmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}_{n\times n} \quad \text{and} \quad A^+A = \begin{bmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}_{m\times m}$$

How to find A^+:

Write $A = U\Sigma V^*$ (the SVD).

Then $\boxed{A^+ = V\Sigma^+U^*}$ where $\Sigma = \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix}$

Note that

$$AA^+ = (U\Sigma V^*)(V\Sigma^+U^*) = U\Sigma\Sigma^+U = \begin{bmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}_{m\times m}$$

$$A^+A = (V\Sigma^+U^*)(U\Sigma V^*) = V\Sigma^+\Sigma V^* = \begin{bmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}_{n\times n}.$$