Section 7: Positive linear maps

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Basic concepts, and relation to eigenvalues

Definition

A self-adjoint map $M\colon X \to X$ is positive-definite (or positive) if

$$(x, Mx) > 0, \qquad \text{for all } x \neq 0,$$

and positive semi-definite (or nonnegative) if

$$(x, Mx) \geq 0, \qquad \text{for all } x \neq 0,$$

We denote these as $M > 0$ and $M \geq 0$, respectively.

Proposition 7.1

A self-adjoint map $M\colon X \to X$ is

(i) positive if and only if all eigenvalues of M are positive,

(ii) non-negative if and only if all eigenvalues of M are nonnegative.

We can define what it means for M to be negative, or non-positive, analogously.

A matrix that is none of these is said to be indefinite.

Basic properties of positive maps

Proposition 7.2

Let X be an inner product space, and $M, N, Q \in \mathrm{Hom}(X, X)$.

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

(ii) If $M > 0$ and Q invertible, then $Q^* M Q > 0$.

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

The topology of positive maps

In an inner product space, the ball of radius $r > 0$ centered at $x \in X$ is

$$B_r(x) = \{y \in X : ||x - y|| < r\}.$$

Let $U \subseteq X$ be a subset. Then

- a point $u \in U$ is interior if there is some $\epsilon > 0$ for which $B_\epsilon(u) \subseteq U$,
- the set U is open if every $u \in U$ is interior,
- its closure consists of U and its limit points.

Proposition 7.3

Let X be an inner product space, and consider the vector space of self-adjoint maps of X.

(i) The subset of positive maps is open.

(ii) The closure of this set are the non-negative maps.

The matrix $A^T A$

Consider an $n \times m$ matrix A over $\mathbb{R}$, where

$$A = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}.$$

The $m \times m$ matrix $A^T A$ is self-adjoint:

$$A^T A = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \cdots & x_1^T x_m \\ x_2^T x_1 & x_2^T x_2 & \cdots & x_2^T x_m \\ \vdots & \vdots & \ddots & \vdots \\ x_m^T x_1 & x_m^T x_2 & \cdots & x_m^T x_m \end{bmatrix}.$$

Note that $A \colon \mathbb{R}^m \to \mathbb{R}^n$ and $A^T A \colon \mathbb{R}^m \to \mathbb{R}^m$. We've already seen that:

1. $\operatorname{rank} A = \operatorname{rank} A^T A$ and $\operatorname{nullity} A = \operatorname{nullity} A^T A$ (in fact, $N_A = N_{A^T A}$),
2. $A^T A \geq 0$, and $A^T A > 0$ if $x_1, \dots, x_m$ are linearly independent,
3. If $N_A = 0$, then the projection matrix onto $\operatorname{Span}(x_1, \dots, x_m)$ is $A(A^T A)^{-1} A^T$.

Later, we'll diagonalize $A^T A$ to get the celebrated singular value decomposition of A.

Gram matrices

Now, we'll generalize the construction of $A^T A$, the "matrix of dot products."

We'll see that *every* positive matrix is a "matrix of inner products."

Definition

Let $x_1, \dots, x_m \in X$, with inner product $(\, , \,)$. The Gram matrix of these vectors is

$$G = (G_{ij}), \qquad \text{where} \quad G_{i,j} = (x_i, x_j).$$

Notice that $G = A^* A$, where $A = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$.

Theorem 7.6

1. Every Gram matrix is nonnegative.
2. The Gram matrix of a set of linearly independent vectors is positive.
3. Every positive matrix is a Gram matrix.

Other examples of Gram matrices

1. Let $X = \{f : [0,1] \to \mathbb{R}\}$, where $(f,g) = \int_0^1 f(t)g(t)\,dt$. If
$$f_1 = 1, \quad f_2 = t, \quad \dots, \quad f_m = t^{m-1},$$
then the Gram matrix is $G = (G_{ij})$, where
$$G_{ij} = \frac{1}{i+j-1}.$$

2. Consider $X = \{f : [0,2\pi] \to \mathbb{C}\}$ and a "weighting function" $w : [0,2\pi] \to \mathbb{R}^+$, define
$$(f,g) = \int_0^{2\pi} f(\theta)\overline{g(\theta)}w(\theta)\,d\theta.$$
If $f_j = e^{ij\theta}$, for $j = -n, \dots, n$, then the $(2n+1) \times (2n+1)$ Gram matrix is $G = (G_{kj}) = (c_{k-j})$, where
$$c_\omega = \int_0^{2\pi} w(\theta)e^{-i\omega\theta}\,d\theta.$$

New inner products from old

Let X be a vector space with inner product $(\cdot\,,\cdot)$.

A positive map $M > 0$ defines a nonstandard inner product $\langle\cdot\,,\cdot\rangle$, where

$$\langle x, y\rangle := (x, My).$$

Lemma 7.7 (HW)

If $H, M \colon X \to X$ are self-adjoint and $M > 0$, then $M^{-1}H$ is self-adjoint with respect to the inner product $\langle x, y\rangle = (x, My)$.

Definition

If $H, M \colon X \to X$ are self-adjoint and $M > 0$, the generalized Rayleigh quotient is

$$R_{H,M}(x) = \frac{(x, Hx)}{(x, Mx)} = \frac{(x, MM^{-1}Hx)}{(x, Mx)} = \frac{\langle x, M^{-1}Hx\rangle}{\langle x, x\rangle} := R_{M^{-1}H}\langle x\rangle \quad \text{w.r.t. } \langle\,,\rangle.$$

Note that:

- the ordinary Rayleigh quotient is simply $R_H = R_{H,I}$.
- the generalized Rayleigh quotient is an ordinary Rayley quotient.

The generalized Rayleigh quotient

Key remark

Results on the generalized Rayleigh quotient $R_{H,M}(x)$ follow from interpreting results of the ordinary Rayleigh quotient to

$$R_{M^{-1}H}\langle x\rangle := \frac{\langle x, M^{-1}Hx\rangle}{\langle x, x\rangle} = \frac{(x, Hx)}{(x, Mx)} = R_{H,M}(x).$$

For example, the minimum value of the Rayleigh quotient is the smallest eigenvalue of H:

$$R_H(v_1) = \lambda_1, \qquad \text{where } Hv_1 = \lambda_1 v_1.$$

The minimum value of the generalized Rayleigh quotient is the smallest eigenvalue of $M^{-1}H$:

$$R_{H,M}(v_1) = R_{M^{-1}H}\langle w_1\rangle = \mu_1, \qquad \text{where } M^{-1}Hw_1 = \mu_1 w_1.$$

Now, w.r.t. the inner product $\langle\, ,\rangle$, let

$$X_1 := \operatorname{Span}(v_1)^{\perp}, \qquad \text{and so} \qquad X = X_1 \oplus \operatorname{Span}(v_1), \qquad \dim X_1 = n-1.$$

The minimum value of the generalized Rayleigh quotient on X_1 is

$$\mu_2 = \min_{||x||=1} \left\{ R_{M^{-1}H}\langle x\rangle \mid \langle x, v_1\rangle = 0 \right\} = \min_{||x|||=1} \left\{ R_{H,M}(x) \mid (x, Mv_1) = 0 \right\}$$

where $M^{-1}Hw_2 = \mu_2 w_2$, and μ_2 is the 2nd smallest eigenvalue of $M^{-1}H$.

The min-max principle for the generalized Rayleigh quotient

Theorem 6.10 (recall)

Let $H\colon X \to X$ be self-adjoint with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$. Then

$$\lambda_k = \min_{\dim S = k} \left\{ \max_{x \in S \setminus 0} R_H(x) \right\}.$$

Proposition 7.8 (HW)

Let $H, M\colon X \to X$ be self-adjoint and $M > 0$.

1. Show that there exists a basis $v_1, \ldots, v_n$ of X where each v_i satisfies

$$Hv_i = \mu_i M v_i \quad (\mu_i \text{ real}), \qquad (v_i, Mv_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

2. Compute (v_i, Hv_j), and show that there is an invertible matrix U for which $U^* M U = I$ and $U^* H U$ is diagonal.
3. Characterize the numbers $\mu_1, \ldots \mu_n$ by a minimax principle.

The Hadamard product of matrices

Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices of the same size. The Hadamard product of A and B is defined as

$$A \circ B := (a_{ij} b_{ij}).$$

Schur's product theorem

If $A, B > 0$, then so is $A \circ B$.

The idea of the polar decomposition

Every nonzero complex number $z \in \mathbb{C}$ has a unique polar form

$$z = re^{i\theta} = |z|e^{i\theta}, \qquad r \in \mathbb{R}^+, \quad \theta \in [0, 2\pi).$$

This can be thought of as decomposing $z \in \mathbb{C}$ into:

- a rotation by θ,
- a scaling by $|z| = r = \sqrt{\bar{z}z}$.

This is simply the polar decomposition of a 1×1 matrix.

Every linear map $A \in \mathrm{Hom}(X, X)$ can be decomposed as $A = UP$, where

- U is unitary; i.e., an isometry of X,
- $P \geq 0$; a scaling along an orthonormal axis $u_1, \ldots, u_n$.

It turns out that $P = \sqrt{A^*A} := |A|$, and so sometimes this is written $A = U|A|$.

In this lecture, we will derive the polar decomposition of a linear map

$$A \colon X \longrightarrow U, \qquad \dim X = m, \quad \dim U = n.$$

In the next lecture, we will derive the celebrated singular value decomposition (SVD).

Singular values

Key properties (Propositions 7.2, 7.6)

- $A^*A \geq 0$;
- Every $P \geq 0$ has a unique nonnegative square root $R := \sqrt{P}$, such that $R^2 = P$.

This means that for some $\lambda_1, \dots, \lambda_m \geq 0$,

$$A^*A = W \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_m^2 \end{bmatrix} W^*, \qquad \text{and} \qquad \sqrt{A^*A} = W \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix} W^*.$$

Definition

The eigenvalues of $\lambda_1, \dots, \lambda_m$ of $\sqrt{A^*A}$ are called the singular values of A.

Facts (that we've seen)

- $||Ax|| = \left|\left|\sqrt{A^*A}x\right|\right|$ for all $x \in X$.
- A, A^*A, and $\sqrt{A^*A}$ have the same nullspace.
- A, A^*A, and $\sqrt{A^*A}$ have the same rank.

Polar decomposition of an invertible map

Theorem

Every linear map $A\colon X \to X$ can be written as $A = UP$ where $P \geq 0$ and U is unitary. This is called the (left) polar decomposition of A.

To construct the polar decomposition, suppose $A = UP$.

Since $P \geq 0$, we can write $P = QDQ^*$,and so

$$P^*P = (QDQ^*)^*(QDQ^*) = (QD^*Q^*)QDQ^* = QD^2Q^* = P^2.$$

Now, notice that

$$A^*A = (UP)^*(UP) = P^*U^*UP = P^*P = P^2.$$

Therefore, $P = \sqrt{A^*A}$.

If A is invertible, then $U = AP^{-1} = A\sqrt{A^*A}^{-1}$ is uniquely determined.

In this case,

$$A = UP = \big(A\sqrt{A^*A}^{-1}\big)\sqrt{A^*A}.$$

If A is not invertible, then U still exists, but is not unique.

Polar decomposition of a general linear map

Theorem

Every linear map $A\colon X \to X$ can be written as $A = UP$ where $P \geq 0$ and U is unitary. This is called the polar decomposition of A.

Suppose the eigenvalues of $\sqrt{A^*A}$ are

$$\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_m = 0,$$

and pick a set $x_1, \dots, x_m$ of orthonormal eigenvectors. Then

$$\frac{1}{\lambda_1}Ax_1, \dots, \frac{1}{\lambda_r}Ax_r, x_{r+1}, \dots, x_m$$

is orthonormal. The polar decomposition is $A = UP$ where $P = \sqrt{A^*A}$ and

$$U = \begin{bmatrix} | & & | & | & & | \\ \frac{1}{\lambda_1}Ax_1 & \cdots & \frac{1}{\lambda_r}Ax_r & x_{r+1} & \cdots & x_m \\ | & & | & | & & | \end{bmatrix} \begin{bmatrix} — & x_1^H & — \\ & \vdots & \\ — & x_m^H & — \end{bmatrix}.$$

Remark

If $A\colon X \to X$ and $r := \det P = |\det A|$, then

$$\det A = \det U \, \det P = e^{i\theta} \cdot r.$$

Singular value decomposition

Need to do. . .

Partially ordered sets

Recall that a partial order on a set X is a relation $\leq$ that is:

(i) reflexive: $x \leq x$

(ii) anti-symmetric: $x \leq y$ and $y \leq x \;\Rightarrow\; x = y$

(iii) transitive: $x \leq y \leq z \;\Rightarrow\; x \leq z$.

We say that $x < y$ if $x \leq y$ and $x \neq y$. The pair $(X, \leq)$ is a partially ordered set (poset).

Alternatively, we can define a partial order by a relation $<$ that is

(i) reflexive: $x \nleq x$

(ii) anti-symmetric: $x < y \;\Rightarrow\; y \nless x$

(iii) transitive: $x < y < z \;\Rightarrow\; x < z$.

Definition

Put a following partial order on the set of self-adjoint maps:

$$M < N \quad \text{iff} \quad N - M > 0, \qquad\qquad M \leq N \quad \text{iff} \quad N - M \geq 0.$$

Basic properties of the poset of positive maps

The following easy facts all hold for positive numbers:

(i) If $m_1 < n_1$ and $m_2 < n_2$, then $m_1 + m_2 < n_1 + n_2$.

(ii) If $\ell < m < n$, then $\ell < n$.

(iii) If $m < n$ and $a > 0$, then $am < an$

(iv) If $0 < m < n$, then $1/m > 1/n > 0$.

Proposition 7.9

The following all hold for linear maps on X:

(i) If $M_1 < N_1$ and $M_2 < N_2$, then $M_1 + M_2 < N_1 + N_2$.

(ii) If $L < M < N$, then $L < N$.

(iii) Given maps $M < N$ and a scalar $a > 0$, we have $aM < aN$.

(iv) If $0 < M < N$, then $M^{-1} > N^{-1} > 0$.

The symmetrized product

Definition

If $A, B \colon X \to X$ are self-adjoint, their symmetrized product is

$$S = AB + BA.$$

Proposition 7.10

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