Lecture 7.2: Nonstandard inner products and Gram matrices

Matthew Macauley

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

Math 8530, Advanced Linear Algebra

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 $x_1, \ldots, x_m$ are linearly independent,
3. If $N_A = 0$, then the projection matrix onto $\operatorname{Span}(x_1, \ldots, x_m)$ 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 $X = \{f \colon [0,1] \to \mathbb{R}\}$, where $(f,g) = \int_0^1 f(t)g(t)\,dt$. If

$$f_1 = 1, \quad f_2 = t, \quad \ldots, \quad f_m = t^{m-1},$$

   then the Gram matrix is $G = (G_{ij})$, where

$$G_{ij} = \frac{1}{i+j-1}.$$

2. Consider $X = \{f \colon [0,2\pi] \to \mathbb{C}\}$ and a "weighting function" $w \colon [0,2\pi] \to \mathbb{R}^+$, define

$$(f,g) = \int_0^{2\pi} f(\theta)\overline{g(\theta)}w(\theta)\,d\theta.$$

   If $f_j = e^{ij\theta}$, for $j = -n, \ldots, n$, 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 7.4 (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 6.8 (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 7.5 (HW)

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

1. Show that there exists a basis $v_1, \dots, v_n$ 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 $\mu_1, \dots \mu_n$ 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 $A \circ B$.