Lecture 6.1: Quadratic forms

Matthew Macauley

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

Math 8530, Advanced Linear Algebra

Self-adjoint and anti-self-adjoint maps

Throughout, let X be a finite-dimensional inner product space.

Definition

A linear map $M\colon X \to X$ is self-adjoint if $M^* = M$, and anti-self-adjoint if $M^* = -M$.

These are also called Hermitian and anti-Hermitian, respectively.

Remark

Every linear map $M\colon X \to X$ can be decomposed into a self-adjoint part and an anti-self-adjoint part:

$$M = H + A, \qquad H = \frac{M + M^*}{2}, \qquad A = \frac{M - M^*}{2}.$$

Compare/contrast this to:

- Every matrix can be written as a sum of a symmetric and skew-symmetric matrix.
- Every real-valued function can be written as a sum of an even and an odd function.

Why do we care about self-adjoint maps?

A real-valued matrix is self-adjoint if it is symmetric: $A^T = A$.

A complex-valued matrix is self-adjoint if it is Hermitian: $\overline{A}^T = A$.

Key idea (preview)

If $A\colon X \to X$ is self-adjoint, then

- all eigenvalues of A are real
- X has an orthonormal basis of eigenvectors of A.

In spaces of functions, self-adjoint differential operators are important because they guarantee an orthogonal basis of eigenfunctions, and a "generalized Fourier series."

Another source of self-adjoint maps are quadratic forms, which we will see in this lecture.

These arise in calculus, statistics, and many other branches of higher mathematics.

We'll begin by motivating them by revisiting the second derivative test from calculus.

Second order approximations

A common problem in Calculus 1 is:

use the tangent line to approximate a function $f(x)$ *near* $a \in \mathbb{R}$.

In Calculus 2, one learns about Taylor series, and higher-order approximations.

For example, consider the function

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots$$

- the 0^{th} order term is $f(a)$
- the 1^{st} order term is $f'(a)$
- the 2^{nd} order term is $\frac{1}{2}f''(a)$.

If a is a critical point (i.e., $f'(a) = 0$), then the behavior of f is governed by $f''(x)$.

Multivariate Taylor series

Now, let $f(x_1, \ldots, x_n)$ be a smooth function $\mathbb{R}^n \to \mathbb{R}$. Then near a point $a \in \mathbb{R}^n$,

$$f(x) = \sum_{k=1}^{\infty} \frac{D^k f(a)}{k!}(x-a)^k = f(a) + \ell(x) + \frac{1}{2}q(x) + \cdots .$$

- the 0^{th} order term is $f(a)$
- the 1^{st} order term is $\ell(y) = \langle g, y\rangle$, where $g = \nabla f(a) = \begin{bmatrix} \frac{\partial f(a)}{\partial x_1} \\ \vdots \\ \frac{\partial f(a)}{\partial x_n} \end{bmatrix}$.
- the 2^{nd} order term is

$$q(y) = \sum_{j=1}^{n}\sum_{i=1}^{n} h_{ij} y_i y_j, \qquad \text{where} \;\; H = (h_{ij}) = \left(\frac{\partial^2 f}{\partial x_i \, \partial x_j} \right)$$

is the Hessian of f. This can be written as

$$q(y) = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{n1} & \cdots & h_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \langle y, Hy \rangle.$$

If $a \in \mathbb{R}^n$ is a critical point (i.e., $\nabla f = 0$), then the behavior of f is governed by $q(y)$.

Quadratic forms

Definition

A quadratic form is a function

$$q\colon X \to K, \qquad q(x) = \langle x, Hx \rangle$$

for some self-adjoint linear map $H\colon X \to X$.

Consider a quadratic form

$$q(x) = x^T Hx = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \langle x, Hx \rangle.$$

If we diagonalize H, i.e., write $H = PDP^{-1} = PDP^T$ (P is orthogonal), then

$$q(x) = \langle x, Hx \rangle = x^T Hx = x^T PDP^T x.$$

If we change variables by letting $z = P^T x$,

$$q(z) = z^T Dz = \begin{bmatrix} z_1 & \cdots & z_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \sum_{i=1}^{n} \lambda_i z_i^2 = \langle z, Dz \rangle.$$

Quadratic forms and conic sections

Consider the quadratic form

$$q(x) = \langle x, Ax \rangle = x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 5x_1^2 - 6x_1x_2 + 5x_2^2.$$

It is easy to check that $A = PDP^T$ (or $D = P^T AP$), where

$$\begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Now, let $z = P^T x$, or $x = Pz$. In this new coordinate system,

$$q(x) = q(Pz) = \langle Pz, APz \rangle = (Pz)^T A(Pz) = z^T P^T APz = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 8z_1^2 + 2z_2^2.$$

Let's sketch the graph of $f(x_1, x_2) = 5x_1^2 - 6x_1x_2 + 5x_2^2 = 1$, which is an ellipse.