Mathematics of Image and Data Analysis
Math 5467

Linear Algebra, Calculus & Python II

Instructor: Jeff Calder
Email: jcalder@umn.edu

http://www-users.math.umn.edu/~jwcalder/5467

Last time

- Projection
- Introduction to Numpy

Today

- Reading images and audio in Python
- Diagonalization
- Some vector calculus

Images and audio in Python (.ipynb)

Diagonalization

Every symmetric matrix A can be diagonalized. That is, there exists an orthogonal matrix Q and a diagonal matrix D such that

$$A = QDQ^T.$$

An orthogonal matrix is a square matrix whose columns are orthonormal vectors.

- The columns of Q are exactly the eigenvectors of the matrix A.
- The diagonal entries of D are the corresponding eigenvalues.
- An orthogonal matrix also has the property that all rows are orthonormal and thus
$$Q^T Q = I = QQ^T.$$
- An orthogonal matrix is norm-preserving
$$\|Qx\| = \|x\|.$$

Optimization and eigenvalues

Exercise 1. Let A be a symmetric matrix, and consider the optimization problem

$$\min\{x^T Ax : \|x\| = 1\}. \tag{1}$$

Show that every minimizer x^* is an eigenvector of A with smallest eigenvalue. What happens if we switch the min to a max in (1)? △

$$A = QDQ^T, \quad x^T A x = x^T Q D Q^T x$$

Set $y = Q^T x$

$$= (Q^T x)^T D (Q^T x)$$

$$= y^T D y$$

$$y^T D y = y^T \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} y = y^T \begin{bmatrix} y(1)\lambda_1 \\ \vdots \\ y(n)\lambda_n \end{bmatrix}$$

$$\|y\| = \|Q^T x\| = \|x\| = 1 \qquad = \sum_{i=1}^{n} y(i)^2 \lambda_i$$

$$\min_{\|y\|=1} \sum_{i=1}^{n} y(i)^2 \lambda_i \ , \quad \lambda_1 \le \cdots \le \lambda_n$$

$$\sum_{i=1}^{n} y(i)^2 \lambda_i \ge \lambda_1 \sum_{i=1}^{n} y(i)^2 = \lambda_1$$

Choosing $y(1) = 1, \ y(2) = 0, \ y(3) = 0, \ \ldots, \ y(n) = 0$

$$\sum_{i=1}^{n} y(i)^2 \lambda_i = \lambda_1 \ , \quad y = e_1$$

is a minimizer

$y = Q^T x$, $\quad x = Qy = Qe_1 = v_1$

v_1 = first eigenvector (smallest eigenvalue λ_1)

$$\lambda_1 < \lambda_2$$

$$\lambda_1 = \lambda_2 = \lambda_3 \cdots = \lambda_k < \lambda_{k+1}$$

Vector Calculus

We recall that for a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, the gradient ∇f is defined by

$$\nabla f = \left(\frac{\partial f}{\partial x(1)}, \frac{\partial f}{\partial x(2)}, \ldots, \frac{\partial f}{\partial x(n)} \right).$$

Example 1. For the function $f(x) = x(1)^2 - x(2)^2$ on $\mathbb{R}^2$, the gradient is

$$\nabla f(x) = (2x(1), -2x(2)).$$

$\triangle$

Gradients of common functions

$f(x) = \sum_{i=1}^{n} y(i)\, x(i)$

Exercise 2. Show that

(i) For a linear function $f(x) = y^T x$, we clearly have $\nabla f(x) = y$.

(ii) For a quadratic function $f(x) = x^T A x$, where A is an $n \times n$ matrix, we have

$$\nabla f(x) = (A + A^T)x.$$

(iii) Assume A is a symmetric matrix. For the function $f(x) = \|Ax\|^2$, show that

$$\nabla f(x) = 2A^2 x,$$

△

$$f(x) = x^T A x = (x^T A x)^T = x^T A^T x$$

$$f(x) = \frac{1}{2} x^T A x + \frac{1}{2} x^T A^T x$$

$$= x^T \left(\underbrace{\tfrac{1}{2}A + \tfrac{1}{2}A^T}_{B} \right) x$$

$B, \quad B = B^T$

$$= x^T B x .$$

We may assume $A = A^T$ (symmetric)

$$\frac{\partial f}{\partial x(k)} = \frac{\partial}{\partial x(k)} \sum_{i=1}^{\wedge} \sum_{j=1}^{\wedge} a(i,j)\, x(i)\, x(j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a(i,j) \frac{\partial}{\partial x(k)} \left(x(i)\, x(j) \right)$$

Kronecker delta

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a(i,j) \left(\delta_{ik}\, x(j) + \delta_{jk}\, x(i) \right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a(i,j)\, x(j)\, \delta_{ik} + \sum_{i=1}^{n} \sum_{j=1}^{n} a(i,j)\, x(i)\, \delta_{jk}$$

$i = k$ $\qquad$ $j = k$

$$= \sum_{j=1}^{n} a(k,j)\,x(j) + \sum_{i=1}^{n} a(i,k)\,x(i)$$

$$= [Ax]_k + [A^T x]_k$$

$$= [(A + A^T)x]_k$$

$$Df = (A + A^T)x$$