Mathematics of Image and Data Analysis
Math 5467

Lecture 3: Linear Algebra & Python

Instructor: Jeff Calder
Email: jcalder@umn.edu

http://www-users.math.umn.edu/~jwcalder/5467S21

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

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

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)? △

Write $y = Q^T x$, $\quad A = Q \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} Q^T$

$\|y\| = \|x\|$

$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$

$x^T A x = x^T Q D Q^T x$

$= y^T \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} y = \sum_{i=1}^{n} \lambda_i y_i^2$

$$\min \left\{ \sum_{i=1}^{n} \lambda_i y_i^2 \;:\; \|y\|^2 = \sum_{i=1}^{n} y_i^2 = 1 \right\}.$$

Minimized by $y_1 = 1, \; y_2 = y_3 = \cdots = y_n = 0$

Indeed, $\displaystyle\sum_{i=1}^{n} \lambda_i y_i^2 \geq \lambda_1 \sum_{i=1}^{n} y_i^2 = \lambda_1$

Thus $y = e_1 = (1, 0, \ldots, 0)$

Check $x = Qy = Qe_1 =$ first eigenvector of A (smallest eigenvalue)

$\rightarrow$ $\boxed{Ax = \lambda_1 x}$

And $x^T A x = x^T \lambda_1 x = \lambda_1 x^T x$

$= \lambda_1 \|x\|^2 = \lambda_1$

$$\boxed{\lambda_1 = x^T A x}$$

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

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

$\triangle$

$x_1, x_2, x_3, x_4 \in \mathbb{R}^n$

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

Gradients of common functions

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 $f(x) = x^T Ax$, 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

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

HW: $0 = \nabla\left(\frac{x^T A x}{x^T x}\right) = \frac{x^T x \, \nabla(x^T A x) - x^T A x \, \nabla(x^T x)}{(x^T x)^2}$

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

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

Can assume A is symmetric and show that for $f(x) = x^T A x$

$$\nabla f(x) = 2Ax$$

$$\frac{\partial}{\partial x(k)} (x^T A x) = \frac{\partial}{\partial x(k)} \sum_{i=1}^{n} \sum_{j=1}^{n} A(i,j)\, x(i)\, x(j)$$

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

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

where

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

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

since $A = A^T$

$$= (Ax)(k) + (Ax)(k)$$

$$= 2\,(Ax)(k)$$

$$\Rightarrow \quad \nabla(x^T A x) = 2Ax$$

for symmetric matrices A

$$A = A^T.$$