Mathematics for Intelligent Systems
Lecture 5 Homework
(Eigenvectors)

Marc Toussaint, Andrea Baisero

1 Problem 1: Eigenvectors

Recall that a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called PSD if $x^\intercal A x \geq 0, \forall x \in \mathbb{R}^n$.

(a) Show that *all* eigenvalues of a positive semidefinite (PSD) matrix are nonnegative.

(b) Show that if v is an eigenvector of A with eigenvalue λ, then v is also an eigenvector of A^k for any positive integer k. What is the corresponding eigenvalue?

(c) A square matrix M is idempotent if $M^2 = M$. What are the possible eigenvalues of an idempotent matrix?

(d) Let v be an eigenvector of A with eigenvalue λ and w an eigenvector of A^T with a different eigenvalue $\mu \neq \lambda$. Show that v and w are orthogonal with respect to the dot product.

(e) Suppose $A \in \mathbb{R}^{n \times n}$ has eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. What are the eigenvalues of $A + \alpha I$ for $\alpha \in \mathbb{R}$ and I an identity matrix?

(f) Assume $A \in \mathbb{R}^{n \times n}$ is diagonalizable, i.e., it has n linearly independent eigenvectors. Initialize $x \in \mathbb{R}^n$ as a random normalized vector and iterate

$$x \leftarrow Ax, x \leftarrow \frac{1}{\|x\|}\, x$$

Prove[1] that under certain conditions $\lambda = x^\intercal A x$ converges to the "largest" (in *absolute* terms $|\lambda_i|$) eigenvalue of A. Discuss these convergence conditions and other convergence properties of this power method.

(g) Let A be a positive definite matrix with λ_{max} its largest eigenvalue (in absolute terms $|\lambda_i|$). What do we get when we apply power iteration method to the matrix $B = A - \lambda_{max}\mathbb{I}$? How can we get the smallest eigenvalue of A?

[1] You may want to refer to Section 4 of Nathan's lecture 3 notes. Se course webpage.

(h) Consider the following variant of the previous power iteration:

$$z \leftarrow Ax, \lambda \leftarrow x^\intercal z, y \leftarrow (\lambda I - A)y, x \leftarrow \frac{1}{\|z\|}\ z, y \leftarrow \frac{1}{\|y\|}\ y$$

If A is a positive definite matrix, show that the algorithm can give an estimate of the "smallest" eigenpair of A.

Exercise 2 - Multivariate Calculus

Given tensors $y \in \mathbb{R}^{a\times\ldots\times z}$ and $x \in \mathbb{R}^{\alpha\times\ldots\times\omega}$ where y is a function of x, the Jacobian tensor of their relation is formally defined as the tensor $J \in \mathbb{R}^{a\times\ldots\times z\times\alpha\times\ldots\times\gamma}$ of partial derivatives

$$J_{i,j,k,\ldots,l,m,n\ldots} = \frac{\partial}{\partial x_{l,m,n,\ldots}} y_{i,j,k,\ldots}$$

(We use the convention that all "output" indices come before all "input" indices".)

Compute the Jacobian tensor for each of the following:

(a) Jacobian of x wrt x.

(b) Jacobian of $c^\intercal x$ wrt x where c is a vector.

(c) Jacobian of $A^\intercal x$ wrt x where A is a vector.

(d) Jacobian of $x^\intercal x$ wrt x.

(e) Jacobian of $x^\intercal Ax$ wrt x where A is a matrix.

(f) Jacobian of $f(x)^\intercal g(x)$ wrt x, where f and g are vector-values functions.