Problem 1.

You're likely familiar with the algebraic definition of the eigenvalues and eigenvectors of a matrix $\mathbf{A}$ as solutions to the equation $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. In this problem, we will prove an additional viewpoint on eigenvalues and eigenvectors that is especially important in network science and many other fields of applied mathematics:

> *For symmetric matrices, eigenvalues/eigenvectors are solutions of optimization problems.*

Let $\|\mathbf{v}\|$ be the Euclidean norm of a vector $\mathbf{v}$. Here's our theorem:

Theorem 1. *Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\mathbf{A}$ ordered from smallest (λ_1) to largest (λ_n). Then,*

$$\lambda_n = \max_{\mathbf{v}\in\mathbb{R}^n} \left\{ \mathbf{v}^T\mathbf{A}\mathbf{v} \mid \|\mathbf{v}\| = 1 \right\} .$$

In words, the largest value of the function $f(\mathbf{v}) = \mathbf{v}^T\mathbf{A}\mathbf{v}$ that can be obtained by picking a vector $\mathbf{v} \in \mathbb{R}^n$ satisfying $\|\mathbf{v}\| = 1$ is λ_n. The reason that Theorem 1 matters from an applied standpoint is that many network science and data analysis tasks correspond to maximizing functions that look like $f(\mathbf{v}) = \mathbf{v}^T\mathbf{A}\mathbf{v}$.

Part (a)

Prove Theorem 1 in three steps.

i. Expand $\mathbf{A}$ in an orthonormal basis of eigenvectors:

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T ,$$

where each $\mathbf{u}_i$ is an eigenvector $\mathbf{A}$ with eigenvalue λ_i. As a reminder, orthonormality means that $\|\mathbf{u}_i\| = 1$ and $\mathbf{u}_i^T\mathbf{u}_j = 1$ if $i = j$ and 0 otherwise. Explain carefully why you are allowed to perform this expansion (cite a theorem).

ii. Write $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$ for some undetermined coefficients $\{\alpha_i\}$, and explain why you are allowed to do this. Compute $\|\mathbf{v}\|$ in terms of the coefficients $\{\alpha_i\}$, and determine what the constraint $\|\mathbf{v}\| = 1$ implies about $\{\alpha_i\}$.

iii. Finally, use the first two steps to compute $\mathbf{v}^T\mathbf{A}\mathbf{v}$ directly in terms of $\{\lambda_i\}$ and $\{\alpha_i\}$. Show that the result is maximized when $\alpha_n = 1$ and $\alpha_i = 0$ for $i < n$. Include a closing sentence to help your reader connect this final result to Theorem 1.

Solution

Part (b)

Prove Theorem 1 (again) using the *principle of Lagrange multipliers*. The principle of Lagrange multipliers states that, if $\mathbf{x}$ solves the problem

$$\max \left\{ f(\mathbf{x}) \mid g(\mathbf{x}) = 0 \right\} \, ,$$

then $\mathbf{x}$ must also satisfy the equation

$$\nabla f(\mathbf{x}) + c \nabla g(\mathbf{x}) = 0$$

for some scalar $c \in \mathbb{R}$, provided that f and g satisfy certain technical conditions (which you may assume are satisfied here).

    You may assume without proof that any continuous function on the unit sphere $S^{n-1} = \{\mathbf{v} : \|\mathbf{v}\| = 1\}$ possesses at least one maximum value (this follows from the compactness of S^{n-1}). You may also assume without proof that the function $f(\mathbf{v}) = \mathbf{v}^T \mathbf{A} \mathbf{v}$ is continuous. State clearly where you use these facts in your proof. Please again make sure to explicitly state where you use the hypothesis that $\mathbf{A}$ is symmetric.

Solution

Part (c)

Using Theorem 1, give an *extremely short* proof that

$$\lambda_1 = \min_{\mathbf{v} \in \mathbb{R}^n} \left\{ \mathbf{v}^T \mathbf{A} \mathbf{v} \mid \|\mathbf{v}\| = 1 \right\} \, .$$

Solution

Problem 2.

Perron-Frobenius problem here.

Solution

Problem 3.

Let $X_1, X_2 \ldots,$ be an infinite sequence of i.i.d. random variables. Let N be a strictly positive random variable that is independent of the sequence $\{X_1, X_2, \ldots\}$. Let

$$Y = \sum_{i=1}^{N} X_i \ .$$

Prove that $\mathbb{E}[Y] = \mathbb{E}[X_1]\mathbb{E}[N]$.

Solution

Problem 4.

In this problem, you will argue that a binomial distribution can be approximated by a Poisson distribution in a certain limit. This approximation will be useful when we study properties of random graph models.

As a reminder, a nonnegative discrete random variable X has a *binomial distribution* with number of trials n and success rate p if its probability mass function is

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \ .$$

A nonnegative discrete random variable Y has a *Poisson distribution* with parameter μ if its probability mass function is

$$p_Y(k) = \frac{\mu^k e^{-\mu}}{k!} \ .$$

Suppose now that X is binomial and that $p = c/n$. Show that, as n grows large, X is approximately Poisson. Identify the parameter of the Poisson distribution in terms of c and n. It suffices to show that the PMF of X converges to a Poisson PMF (formally, we are showing *convergence in distribution*).

The following approximations are valid for large n, when k and m are fixed constants. You may use them without proof.

$$\binom{n}{k} \simeq \frac{n^k}{k!} \quad \text{and} \quad e^m \simeq \left(1 + \frac{m}{n}\right)^n \ .$$

Solution

Problem 5.

Please briefly respond to the following questions:

    i. Why are you interested in learning about networks?

    ii. What is one question you have about the role that networks play in your life today?

    iii. What is something related to networks that you might like to achieve by the end of the quarter?

    iv. MATH 168 has upper-div prerequisites in linear algebra and probability. Pick one of these two prerequisites. Share something that you thought was especially cool or interesting, and something that you found especially disappointing or frustrating.

A few sentences for each item is plenty.