Special Topics in Applied Mathematics I
Solution 1

XIA JIAHAN

January 21, 2026

Question 1

Iterative eigenvalue methods: an intuitive story

Think of a matrix A as a machine that acts on vectors: it can rotate, shear, and stretch them. Most directions are changed in complicated ways, but some special directions come back pointing the same way after applying A. Those directions are *eigenvectors* $v \neq 0$ satisfying

$$Av = \lambda v,$$

where λ is the corresponding *eigenvalue* (the factor by which A stretches v). When matrices are large, computing all eigenvalues exactly (e.g., from a characteristic polynomial) is impractical and numerically fragile. Iterative methods instead exploit one operation we can often do efficiently: repeatedly multiply by A (or solve linear systems involving A).

Power iteration: repeated multiplication reveals the dominant direction

The simplest method is *power iteration*. Choose any nonzero starting vector b_0 and repeat:

$$c_{k+1} = Ab_k,$$
$$b_{k+1} = \frac{c_{k+1}}{\|c_{k+1}\|}.$$

Normalization "resets the length" so we can watch the direction. The method tends to align b_k with the eigenvector associated with the eigenvalue of largest magnitude.

Why it converges (qualitative). Assume A is diagonalizable with eigenpairs (λ_i, v_i) and $|\lambda_1| > |\lambda_2| \geq \cdots$. Write the starting vector as a combination of eigenvectors:

$$b_0 = \sum_{i=1}^{n} \alpha_i v_i \quad (\alpha_1 \neq 0).$$

Then

$$A^k b_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k v_i = \lambda_1^k \left(\alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \Big(\frac{\lambda_i}{\lambda_1}\Big)^k v_i \right).$$

As k grows, the ratios $\left(\frac{\lambda_i}{\lambda_1}\right)^k$ shrink toward 0, so the direction of $A^k b_0$ becomes dominated by v_1. The convergence speed is governed by the spectral gap; a common rule of thumb is that the directional error shrinks like

$$\left|\frac{\lambda_2}{\lambda_1}\right|^k.$$

Estimating the eigenvalue: Rayleigh quotient

Once we have an approximate eigenvector x, a natural estimate of its eigenvalue is the *Rayleigh quotient*

$$R(x) = \frac{x^\top A x}{x^\top x}.$$

If x is exactly an eigenvector v, then $R(v) = \lambda$ because $v^\top A v = v^\top(\lambda v) = \lambda\, v^\top v$. When x is close to an eigenvector, $R(x)$ typically gives a good scalar approximation of the associated eigenvalue, so one can monitor the sequence $R(b_k)$ during power iteration.

Targeting an eigenvalue near a chosen number: shift-and-invert ("closest to π")

Power iteration finds the eigenvalue with largest magnitude, but sometimes we want the eigenvalue *closest* to a target μ (here $\mu = \pi$). The key idea is to transform the problem so that "closest to μ" becomes "dominant in magnitude".

Shifted inverse iteration. Consider the matrix $(A-\mu I)^{-1}$ (assuming $A-\mu I$ is invertible). If $Av = \lambda v$, then

$$(A - \mu I)^{-1} v = \frac{1}{\lambda - \mu}\, v.$$

So the eigenvalues of $(A - \mu I)^{-1}$ are $\frac{1}{\lambda_i - \mu}$, with the same eigenvectors as A. The eigenvalue of A closest to μ makes $|\lambda_i - \mu|$ smallest, hence $\left|\frac{1}{\lambda_i - \mu}\right|$ largest, becoming dominant for the transformed matrix.

The iteration is therefore:

$$\begin{aligned} (A - \mu I) y_{k+1} &= b_k, \\ b_{k+1} &= \frac{y_{k+1}}{\|y_{k+1}\|}, \\ \lambda_{k+1} &= R(b_{k+1}) \quad \text{(optional eigenvalue estimate).} \end{aligned}$$

This is power iteration applied to $(A - \mu I)^{-1}$, implemented by solving a linear system at each step.

Demonstration: convergence to the eigenvalue closest to π. Take a toy example where we can see everything clearly:

$$A = \mathrm{diag}(3.14,\ 1,\ -2), \qquad \mu = \pi \approx 3.14159.$$

The eigenvalues of A are 3.14, 1, and -2, and the one closest to π is 3.14. For the shifted inverse map,

$$(A - \pi I)^{-1} = \mathrm{diag}\left(\frac{1}{3.14 - \pi},\ \frac{1}{1 - \pi},\ \frac{1}{-2 - \pi}\right),$$

so the magnitudes are approximately

$$\left|\frac{1}{3.14 - \pi}\right| \approx 628, \qquad \left|\frac{1}{1 - \pi}\right| \approx 0.466, \qquad \left|\frac{1}{-2 - \pi}\right| \approx 0.241.$$

The dominant term corresponds to the eigenvalue 3.14, because it is closest to π. Starting from almost any b_0 with a nonzero first component, repeated shifted inverse iteration amplifies the first coordinate relative to the others; after normalization, b_k aligns with the eigenvector e_1, and the Rayleigh quotient $R(b_k)$ converges to 3.14. This illustrates the general fact: shift-and-invert targets

$$\arg\min_i |\lambda_i - \mu|,$$

so choosing $\mu = \pi$ drives the method toward the eigenvalue of A nearest π.