4. By the power method,
\[ v_k = \frac{1}{c} (\lambda_1^k a_1 + \ldots + \lambda_n^k a_n), \]
where \( C = (\lambda_1^2 a_1^2 + \ldots + \lambda_n^2 a_n^2) \), a normalizing constant such that \( \|v_k\|_2 = 1 \). Then
\[ v_k^T A v_k = v_k^T (\lambda_1^{k+1} a_1 + \ldots + \lambda_n^{k+1} a_n) \]
\[ = \frac{1}{c^2} (\lambda_1^{2k+1} a_1^2 + \ldots + \lambda_n^{2k+1} a_n^2) \]
\[ = \lambda_1^{2k+1} a_1^2 + \ldots + \lambda_n^{2k+1} a_n^2 \]
\[ = \frac{1}{1 + (\frac{\lambda}{\lambda_1})^{2k}} (\frac{\lambda}{\lambda_1})^{2k} \lambda_1 \frac{(\frac{\lambda}{\lambda_1})^2}{\lambda_1} \]
\[ = \lambda_1 + O(|\lambda_2/\lambda_1|^{2k}) \]
\[ = (\lambda_1 + O(|\lambda_2/\lambda_1|^{2k})) (1 + O(|\lambda_2/\lambda_1|^{2k})) \]
\[ = \lambda_1 + O(|\lambda_2/\lambda_1|^{2k}) \]
Recall that as in perturbation analysis, since \( \frac{1}{1+x} = 1 - x + x^2 \ldots \), we have \( \frac{1}{1+O(x)} = 1+O(x) \) as \( x \to 0 \)

5. \( A - \mu I \) has eigenvalues \( \lambda_1 - \mu > \lambda_2 \geq \ldots \geq \lambda_{n-1} - \mu > \lambda_n - \mu \). To attain the fastest convergence to the eigenvector corresponding to \( \lambda_1 \), we need
\[ |\lambda_1 - \mu| = \max_{1 \leq i \leq n} |\lambda_i - \mu| = \max(|\lambda_1 - \mu|, |\lambda_n - \mu|), \]
and
\[ \mu = \arg \min \left( \max_{1 \leq i \leq n} |\lambda_i - \mu| \right) = \arg \min \left( \max(|\lambda_2 - \mu|, |\lambda_n - \mu|) \right) \]
The second equation holds when
\[ |\lambda_2 - \mu| = |\lambda_n - \mu| \]
Therefore
\[ \mu = \frac{\lambda_2 + \lambda_n}{2} \]
To achieve the fastest convergence to \( \lambda_n \), we need
\[ |\lambda_n - \mu| = \max_{1 \leq i \leq n} |\lambda_i - \mu| = \max(|\lambda_1 - \mu|, |\lambda_n - \mu|), \]
and
\[ \mu = \arg \min \left( \max_{1 \leq i \leq n-1} |\lambda_i - \mu| \right) = \arg \min \left( \max(|\lambda_1 - \mu|, |\lambda_{n-1} - \mu|) \right) \]
The second equation holds when
\[ |\lambda_1 - \mu| = |\lambda_{n-1} - \mu| \]
Therefore
\[ \mu = \frac{\lambda_1 + \lambda_{n-1}}{2} \]