Power Method

Let $\underline{\underline{A}} \in \mathbb{R}^{n \times n}_{sym}$ and Consider the iteration:

$$\begin{cases} \underline{w}_{k+1} = \underline{\underline{A}}\, \underline{v}_k \\ \underline{v}_{k+1} = \underline{w}_{k+1} / \|\underline{w}_{k+1}\| \end{cases} \qquad (\underline{v}_0 \in \mathbb{R}^n \text{ given})$$

Claim: $\underline{x}_k \to \underline{x}$, the eigenvector with maximal eigenvalue:

$$\underline{\underline{A}}\, \underline{x} = \lambda_{max}(\underline{\underline{A}})\, \underline{x},$$

$$\lambda_{max}(\underline{\underline{A}}) \triangleq \max_{i=1} |\lambda_i(\underline{\underline{A}})| \quad (\text{Denote by } k^*)$$

Pf: We decompose

(sketch)

$$\underline{v}_k = \sum_{i=1}^{n} \alpha_i \underline{x}^{(i)}, \qquad \alpha_i = \underline{v}_k^T \underline{x}^{(i)}.$$

$$\Rightarrow \underline{w}_{k+1} = \sum_{i=1}^{n} \alpha_i \lambda_i \underline{x}^{(i)}$$

$$\Rightarrow \|\underline{w}_{k+1}\| = \left(\sum_{i=1}^{n} \alpha_i^2 \lambda_i^2\right)^{1/2}$$

$$\Rightarrow \underline{v}_{k+1} = \sum_{i=1}^{n} \alpha_i \lambda_i \underline{x}^{i} \Big/ \left(\sum_{i=1}^{n} \alpha_i^2 \lambda_i^2\right)^{1/2}$$

$$= \frac{\alpha_{k^*} \lambda_{k^*} \underline{x}^{(k^*)} + \sum_{i \neq k^*} \alpha_i \lambda_i \underline{x}^{(i)}}{\left(\alpha_{k^*}^2 \lambda_{k^*}^2 + \sum_{i \neq k^*} \alpha_i^2 \lambda_i^2\right)^{1/2}}$$

$| \; | < 1!$

$$= \frac{\underline{x}^{(k^*)} + \sum_{i \neq k^*} (\alpha_i/\alpha_{k^*}) \cdot \left(\frac{\lambda_i}{\lambda_{k^*}}\right) \underline{x}^{(i)}}{\left(1 + \sum_{i \neq k^*} (\alpha_i/\alpha_{k^*})^2 (\lambda_i/\lambda_{k^*})^2\right)^{1/2}}$$

So that, after m steps

$$\underline{v}^{(k+m)} = \frac{\left(\underline{x}^{(k^*)} + \sum_{i \neq k^*} (\alpha_i / r_{k^*}) \cdot \left(\frac{\lambda_i}{\lambda_{k^*}}\right)^m \underline{x}^{(i)}\right)}{c}$$

where c ensures that $\| \underline{v}^{(k+m)} \| = 1$.

Because $| \lambda_i / \lambda_{k^*} | < 1$ by definition of k^*, all that is left in the limit as $m \to \infty$ is the term proportional to $\underline{x}^{(k^*)}$! $\square$

Inverse Iteration

Say we have a good approximation ϑ of some λ for a matrix $\underline{\underline{A}} \in \mathbb{R}^{n\times n}_{sym}$ (e.g. via Gershgorin or other means).

$\underline{Q}$: How do we find $\underline{v}$ s.t. $\underline{\underline{A}}\,\underline{v} = \lambda \underline{v}$? Can we refine ϑ to a better estimate of λ?

<u>Alg</u> (Inverse Iteration)

Input: $\vartheta \in \mathbb{R}$, $\vartheta \approx \lambda$, $\underline{v}^{(0)} \in \mathbb{R}^n$, $\|\underline{v}^{(0)}\| = 1$, $\underline{v}^{(0)} \approx \underline{v}$.

Define: $$(\underline{\underline{A}} - \vartheta \underline{\underline{I}})\,\underline{w}^{(k)} = \underline{v}^{k}$$

$$\underline{v}^{(k+1)} = \underline{w}^{(k)} / \|\underline{w}^{(k)}\|$$

<u>Thm</u>: (Convergence of the Inverse Iteration)

Let $\underline{\underline{A}} \in \mathbb{R}^{n\times n}_{sym}$. The sequence of vectors $\underline{v}^{(k)}$ converges to $\underline{v} \in \mathbb{R}^n$ with $\|\underline{v}\| = 1$, $\underline{\underline{A}}\,\underline{v} = \lambda \underline{v}$ as long as $\underline{v}^{(0)} \cdot \underline{v} \neq 0$. Here, λ is the closest eigenvalue of $\underline{\underline{A}}$ to ϑ.

<u>Pf</u>: By properties of symmetric matrices,

$$\underline{v}^{(0)} = \sum_{i=1}^{n} \gamma_i \underline{x}^{(i)}, \quad \gamma_i = (\underline{x}^{(i)})^T \underline{v}^{(0)}$$

Let $\lambda_s = \lambda$. We want to show that

$$\lim_{k\to\infty} \underline{v}^{(k)} = \underline{x}^{(s)}.$$

if $\gamma_s = (\underline{x}^{(s)})^T \underline{v}^{(0)} \neq 0$. Let us expand

$$w^{(0)} = \sum^{n} \beta_i x^{(i)}$$

$$\Rightarrow (\underline{\underline{A}} - \vartheta \underline{\underline{I}}) \underline{w}^{(0)} = \sum_{i=1}^{n} \beta_i (\lambda_i - \vartheta) \underline{x}^{(i)} = \sum_{i=1}^{n} \gamma_i \underline{x}^{(i)}$$

$$\Rightarrow \gamma_i = \beta_i (\lambda_i - \vartheta) \quad \text{(by orthogonality)}$$

Because $\gamma_s \neq 0$, we have $\lambda_s \neq \vartheta$ (otherwise $\beta_s = \infty$)

Then $\lambda_i - \vartheta \neq 0$ either, because ϑ closest to λ_s.

$$\Rightarrow \underline{v}^{(1)} = C_0 \cdot \sum_{i=1}^{n} \left(\frac{\gamma_i}{\lambda_i - \vartheta} \right) \underline{x}^{(i)}$$

$$\vdots$$

$$\Rightarrow \underline{v}^{(m)} = C_{m-1} \cdots C_0 \sum_{i=1}^{n} \left(\frac{\gamma_i}{(\lambda_i - \vartheta)^m} \right) \underline{x}^{(i)}$$

To ensure that $\| \underline{v}^{(m)} \| = 1$, we have that

$$C_{m-1} \cdots C_0 = \left[\sum_{i=1}^{n} \gamma_i^2 / (\lambda_i - \vartheta)^{2m} \right]^{-1/2}$$

So that

$$\underline{v}^{(m)} = \sum_{i=1}^{n} \left(\frac{\gamma_i}{(\lambda_i - \vartheta)^m} \right) \underline{x}^{(i)} \Big/ \left[\sum_{i=1}^{n} \gamma_i^2 / (\lambda_i - \vartheta)^{2m} \right]^{1/2}$$

Let us write:

$$\left[\sum_{i=1}^{n} \gamma_i^2 / (\lambda_i - \vartheta)^{2m} \right]^{1/2} = \left(\frac{\gamma_s^2}{(\lambda_s - \vartheta)^{2m}} + \sum_{i \neq s} \frac{\gamma_i^2}{(\lambda_i - \vartheta)^{2m}} \right)^{1/2}$$

$$= \left(\frac{\gamma_s}{(\lambda_s - \vartheta)^m} \right) \left(1 + \sum_{i \neq s} \left(\frac{\gamma_i^2}{\gamma_s^2} \right) \left(\frac{\lambda_s - \vartheta}{\lambda_i - \vartheta} \right)^{2m} \right)^{1/2}$$

$$\Rightarrow \underline{v}^{(m)} = \frac{\frac{\alpha_s}{(\lambda_s - \sigma)^m} \underline{x}_s + \sum_{i \neq s}^{n} \left(\frac{\alpha_i}{(\lambda_i - \sigma)^m} \right) \underline{x}^{(i)}}{\left(\frac{\alpha_s}{(\lambda_s - \sigma)^m} \right) \left(1 + \sum_{i \neq s} \left(\frac{\alpha_i^2}{\alpha_s^2} \right) \left(\frac{\lambda_s - \sigma}{\lambda_i - \sigma} \right)^{2m} \right)^{1/2}}$$

$$= \left[\frac{\underline{x}_s + \sum_{i \neq s}^{n} \left(\frac{\alpha_i}{\alpha_s} \right) \left(\frac{\lambda_s - \sigma}{\lambda_i - \sigma} \right)^m \underline{x}^{(i)}}{\left(1 + \sum_{i \neq s} \left(\frac{\alpha_i^2}{\alpha_s^2} \right) \left(\frac{\lambda_s - \sigma}{\lambda_i - \sigma} \right)^{2m} \right)^{1/2}} \right]$$

Now, by definition, $\left| \lambda_s - \sigma / \lambda_i - \sigma \right| < 1$, so that $\underline{v}^{(m)} \longrightarrow \underline{x}^{(s)}$ as $m \to \infty$. $\square$

Note: Proof breaks down if $\alpha_s = 0$, but in practice rounding errors ensure that $\alpha_s \neq 0$ exactly.

Note: Problem if there is a multiple eigenvalue, or two close together.

Rayleigh Quotient

Q: Given an estimate of $\underline{v}$, can we find λ?

Def: Given $\underline{x} \in \mathbb{R}^n$, and an $\underline{\underline{A}} \in \mathbb{R}^{n \times n}_{sym}$, the Rayleigh Quotient is defined as the number

$$R(x) = \frac{\underline{x}^T \underline{\underline{A}}\, \underline{x}}{}$$

$$\underline{x}^T \underline{x}$$

Note: If $\underline{\underline{A}}\,\underline{x} = \lambda \underline{x}$, $R(\underline{x}) = \lambda$!

Note: By expansion and orthonormality,

$$\underline{x} = \sum_{j=1}^{n} r_j \underline{x}^{(j)},$$

$$\underline{\underline{A}}\,\underline{x} = \sum_{j=1}^{n} r_j \lambda_j \underline{x}^{(j)},$$

$$\underline{x}^T \underline{\underline{A}}\,\underline{x} = \sum_{j=1}^{n} r_j^2 \lambda_j$$

$$\underline{x}^T \underline{x} = \sum_{j=1}^{n} r_j^2$$

$$\therefore\ R(\underline{x}) = \sum_{j=1}^{n} r_j^2 \lambda_j \Big/ \sum_{j=1}^{n} r_j^2$$

This leads to the easy Corollary:

Thm: $\lambda_{min}(\underline{\underline{A}}) \le R(x) \le \lambda_{max}(A)$

Pf:

$$\lambda_{min}\left(\sum_j r_j^2 \Big/ \sum_j r_j^2\right) \le \sum_j^n r_j^2 \lambda_j \Big/ \sum_{j=1}^n r_j^2 \le \lambda_{max} \cdot \left(\sum_j r_j^2 \Big/ \sum_j r_j^2\right)$$

$\square$

The following theorem gives us a bound on the accuracy of the eigenvalue, given accuracy on the eigenvector.

Thm: Let $\underline{x} \in \mathbb{R}^n$ be such that

$$\|\underline{x}\| = 1, \quad \|\underline{x} - \underline{x}^{(k)}\| = O(\varepsilon), \quad \varepsilon \ll 1.$$

Then, $R(\underline{x}) = \lambda_k + O(\varepsilon^2)$.

Pf: Because $\underline{x}^T \underline{x}^{(k)} = r_k$ by def,

$$\|\underline{x} - \underline{x}^{(k)}\|^2 = \|x\|^2 - 2\underline{x}^T \underline{x}^{(k)} + \|\underline{x}^{(k)}\|^2$$

$$= 1 - 2r_k + 1$$

$$= 2(1-\tau_K)$$

$$\Rightarrow \tau_K = 1 + O(\varepsilon^2).$$

Hence,

$$\|\underline{x}\|^2 = 1 = \sum_{j=1}^{n} \tau_j^2$$

$$= \tau_K^2 + \sum_{j \neq K}^{n} \tau_j^2$$

$$= \left(1 + O(\varepsilon^2)\right)^2 + \sum_{j \neq K}^{n} \tau_j^2$$

$$\Rightarrow \quad 1 = 1 + O(\varepsilon^2) + \sum_{j \neq K}^{n} \tau_j^2$$

$$\Rightarrow \sum_{j \neq K}^{n} \tau_j^2 = O(\varepsilon^2)$$

So that $\tau_j = O(\varepsilon^2)$ for all $j \neq K$. Now,

$$R(x) = \left(\lambda_K \tau_K^2 + \sum_{j \neq K} \lambda_j \tau_j^2\right) \Big/ \underbrace{\|x\|^2}_{=1}$$

$$= \lambda_K(1 + O(\varepsilon^2)) + O(\varepsilon^2)$$

$$= \lambda_K + O(\varepsilon^2) \qquad \square$$

So the Rayleigh quotient improves over the accuracy of the eigenvector by one power of ε.