Recall that:

Thm: Let $\underline{\underline{A}} \in \mathbb{R}^{n\times n}_{sym.}$ Then,

(i) $\exists$ n linearly independent eigenvectors $\underline{x}^{(i)}$ with

$$\underline{\underline{A}}\,\underline{x}^{(i)} = \lambda \underline{x}^{(i)}$$

(ii) The function $\lambda \mapsto \det(\underline{\underline{A}} - \lambda \underline{\underline{I}})$ is called the characteristic polynomial of $\underline{\underline{A}}$. The eigenvalues are the zeros of this polynomial.

(iii) If $\lambda^{(i)} \neq \lambda^{(j)}$, $\underline{x}^{(i)} \cdot \underline{x}^{(j)} = 0$ (orthogonality)

(iv) If $\underline{\underline{B}} = \underline{\underline{\Phi}}^T \underline{\underline{A}}\,\underline{\underline{\Phi}}$ with $\underline{\underline{\Phi}}$ orthogonal, then the eigenvalues of B are the same as those of $\underline{\underline{A}}$, and the eigenvectors are $\underline{\underline{\Phi}}^T \underline{x}^{(i)}$.

(v) If λ_i has multiplicity m, then there is a linear subspace in $\mathbb{R}^n$ of dimension m spanned by m mutually orthogonal eigenvectors associated to λ_i.

(vi) If the eigenvectors are normalized ($\|x^{i}\| = 1$), let $\underline{\underline{X}}$ have columns the $\underline{x}^{(i)}$. Then

$$\underline{\underline{\Lambda}} = \underline{\underline{\operatorname{diag}}}(\{\lambda^{(i)}\}) = \underline{\underline{X}}^T \underline{\underline{A}}\,\underline{\underline{X}}$$

(vii) Any $\underline{v} \in \mathbb{R}^n$ may be written

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

(viii) $\operatorname{Tr}(\underline{\underline{A}}) = \sum_{i=1}^{n} \lambda_i$.

Gershgorin Theorems

Q: Can we estimate the locations of the λ_i for $\underline{\underline{A}} \in \mathbb{C}^{n\times n}$?

Def: Let $\underline{\underline{A}} \in \mathbb{C}^{n\times n}$, $n \geq 2$. The Gershgorin Discs D_i, $i=1,\dots,n$ of the matrix $\underline{\underline{A}}$ are the closed circular regions

$$D_i = \{ z \in \mathbb{C} \mid |z - a_{ii}| \leq R_i \} \subset \mathbb{C}$$

with $R_i = \sum_{j \neq i} |a_{ij}|$.

Thm: Let $n \geq 2$ and $\underline{\underline{A}} \in \mathbb{C}^{m\times n}$. All eigenvalues of the matrix $\underline{\underline{A}}$ lie in the region $D \triangleq \bigcup_{i=1}^{n} D_i$.

Pf: Let $\lambda \in \mathbb{C}$, $\underline{x} \in \mathbb{C}^n$, $\underline{x} \neq 0$ be an eigenvector/value pair. Then,

$$\sum_{j=1}^{n} a_{ij} x_j = \lambda x_i$$

Let $k = \underset{j=1,\dots,n}{\operatorname{argmax}} |x_j|$. $\Rightarrow |x_j| \leq |x_k|$ for all j. Then

$$|\lambda - a_{kk}|\,|x_k| = |\lambda x_k - a_{kk} x_k|$$

$$= \left| \sum_{j=1}^{n} a_{kj} x_j - a_{kk} x_k \right| \quad (\underline{\underline{A}}\,\underline{x} = \lambda \underline{x})$$

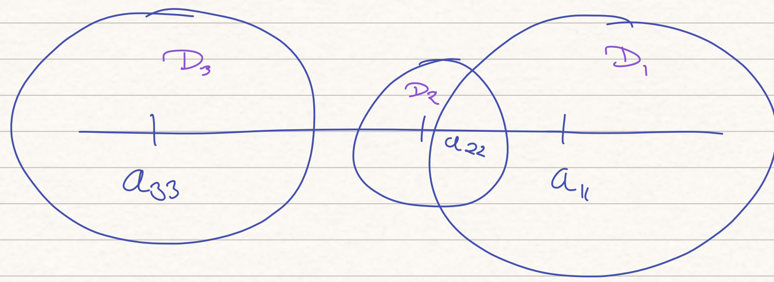
$$= \left| \sum_{j \neq k} a_{kj} x_j \right|$$

$$\leq R_k \cdot |x_k|$$

$$\Rightarrow |\lambda - a_{kk}| \leq R_k$$

$$\Rightarrow \lambda \in D_k \subseteq \bigcup_{i=1}^{n} D_i \quad \square$$

Ex: $\underset{=}{A} = \begin{pmatrix} 3 & 1 & -1/2 \\ 1 & 2 & 0 \\ 1 & 1/2 & -1 \end{pmatrix}$ $\quad R_1 = 3/2 \quad R_2 = 1 \quad R_3 = 3/2$



Thm (Gershgorin's 2nd Thm):

Let $n \geq 2$. Suppose $1 \leq p \leq n-1$. Assume the Gershgorin discs can be divided into two disjoint subsets $D^{(p)}$ and $D^{(q)}$ containing p and $q = n-p$ discs, respectively. Then the union of the discs in $D^{(p)}$ contains exactly p eigenvalues, and the union of the discs in $D^{(q)}$ exactly q. In particular, if one disc is disjoint, then it contains exactly one eigenvalue.

Pf: Consider the matrix

$$b_{ij}(\varepsilon) = \begin{cases} a_{ii} & i = j \\ \varepsilon a_{ij} & i \neq j \end{cases}$$

Note: $\underset{=}{B}(0) = \operatorname{diag}(\underset{=}{A})$, $\quad B(1) = \underset{=}{A}$.

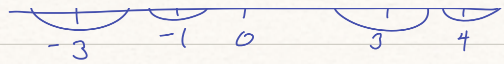
Eigenvalues of $\underset{=}{B}(0)$ are centers of the D_i!

So p of the eigenvalues of $\underset{=}{B}(0)$ lie in the discs inside $D^{(p)}$.

By continuity arguments in ε, must also be true for $B(\varepsilon)$ w/ $\varepsilon \in [0, 1]$, and hence true for $\underset{=}{A}$ (see book for details) $\square$

Ex: $\begin{pmatrix} 4 & 0.2 & -0.1 & 0.1 \\ 0.2 & -1 & -0.1 & 0.05 \end{pmatrix}$

$$\underline{\underline{A}} = \begin{pmatrix} -0.1 & -0.1 & 3 & 0.1 \\ 0.1 & 0.05 & 0.1 & -3 \end{pmatrix}$$



Disjoint discs give bounds on the eigenvalues!

$\lambda_1 \in [3.6, 4.4]$ $\qquad$ $\lambda_3 \in [2.7, 3.3]$

$\lambda_2 \in [-1.35, -0.65]$ $\qquad$ $\lambda_4 \in [-3.25, -2.75]$ $\quad \Delta$

Power Method / 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).

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?

Alg (Inverse Iteration)

Input: $\vartheta \in \mathbb{R}$, $\vartheta \simeq \lambda$, $\underline{v}^{(0)} \in \mathbb{R}^n$, $\|\underline{v}^{(0)}\| = 1$, $\underline{v}^{(0)} \not\simeq \underline{v}$.

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

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

Thm: (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.

Pf: 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

$$\underline{w}^{(0)} = \sum_{i=1}^{n} \beta_i \underline{x}^{(i)}$$

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

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

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

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

Let us write:

$$\left[\sum_{j=1}^{n} \alpha_j / (\lambda_j - \vartheta)^{2m}\right] = \left(\frac{\alpha_s}{(\lambda_s - \vartheta)^{2m}} + \sum_{j \neq s} \frac{\alpha_j}{(\lambda_j - \vartheta)^{2m}}\right)$$

$$= \left(\frac{\alpha_s}{(\lambda_s - \vartheta)^m}\right)\left(1 + \sum_{j \neq s} \left(\frac{\alpha_j^2}{\alpha_s^2}\right)\left(\frac{\lambda_s - \vartheta}{\lambda_j - \vartheta}\right)^{2m}\right)^{1/2}$$

$$\Rightarrow \underline{v}^{(m)} = \frac{\dfrac{\alpha_s}{(\lambda_s - \vartheta)^m} X_s + \displaystyle\sum_{j \neq s}^{n} \left(\frac{\alpha_j}{(\lambda_j - \vartheta)^m}\right) \underline{X}^{(j)}}{\left(\dfrac{\alpha_s}{(\lambda_s - \vartheta)^m}\right)\left(1 + \displaystyle\sum_{j \neq s} \left(\frac{\alpha_j^2}{\alpha_s^2}\right)\left(\frac{\lambda_s - \vartheta}{\lambda_j - \vartheta}\right)^{2m}\right)^{1/2}}$$

$$= \left[\frac{X_s + \displaystyle\sum_{j \neq s}^{n} \left(\frac{\alpha_j}{\alpha_s}\right)\left(\frac{\lambda_s - \vartheta}{\lambda_j - \vartheta}\right)^{m} X^{(j)}}{\left(1 + \displaystyle\sum_{j \neq s} \left(\frac{\alpha_j^2}{\alpha_s^2}\right)\left(\frac{\lambda_s - \vartheta}{\lambda_j - \vartheta}\right)^{2m}\right)^{1/2}}\right]$$

Now, by definition, $|\lambda_s - \vartheta / \lambda_j - \vartheta| < 1$, so that $\underline{v}^{(m)} \longrightarrow \underline{X}^{(s)}$ as $m \to \infty$. $\square$