Lenny Wu

Algorithm 1. Power Method

- Choose $\vec{x}^{(0)} \neq 0$ such that $a_1 \neq 0$ in the representation
  ($\vec{x}^{(0)} = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots a_n\vec{v}_n$, where $\vec{v}_i$ are
  eigenvectors corresponding to λ_i)
- For $k = 1, 2, \dots$

  $\vec{x}^{(k)} = A\vec{x}^{(k-1)}$ (eigenvectors)

  $$\mu_k = \frac{\vec{x}^{(k)T} A \vec{x}^{(k)}}{\vec{x}^{(k)T}\vec{x}^{(k)}}$$

If $\|\mu^{(k)} - \mu^{(k-1)}\| <$ tol., then stop and return $x^{(k)}$ as the appropriate dominant eigenvector corresponding to dominant eigenvalue $\mu^{(k)}$.

Rmk. Drawback of this version is that
$\lim_{k\to\infty} \lambda_1^k$ causes numerical instabilities.
$\therefore$ we consider scaling each $\vec{x}^{(k)}$.

Algorithm 2. Power Method with Scaling

- Choose $\vec{x}^{(0)} \neq 0$ such that $\|\vec{x}^{(0)}\| = 1$ and $a_1 \neq 0$.
- For $k = 1, 2, \dots$

$$\vec{w}^{(k)} = A\vec{x}^{(k-1)}$$

$$\vec{x}^{(k)} = \frac{\vec{w}^{(k)}}{\|\vec{w}^{(k)}\|} \qquad \text{normalize}$$

$$\mu^{(k)} = \vec{x}^{(k)T} A \vec{x}^{(k)} \qquad \text{Rayleigh Quotient}$$

If $|\mu^{(k)} - \mu^{(k-1)}| \leq$ tol. or $\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\| <$ tol.,
then stop, return $\vec{x}^{(k)}$ and $\mu^{(k)}$.

Rmk. 1. If the ∞-norm is used, then it yields Alg. 9.1 in the book. It converges linearly to λ_1 with $O(|\frac{\lambda_2}{\lambda_1}|^k)$.

2. If the l_2-norm is used, then it yields Alg. 9.2 in the book and works well for symmetric matrices with $|\mu^{(k)} - \lambda_1| = O(|\frac{\lambda_2}{\lambda_1}|^{2k})$.

3. Power method is limited by
   - linear convergence
   - if $|\lambda_1| \approx |\lambda_2|$, it is very slow

4. Linear convergence $\Rightarrow$ Aitken's $\{\Delta^2\}$ method can be applied to speed up convergence for estimating dominant eigenvalue.

$$\hat{\mu}^{(k+1)} = \hat{\mu}^{(k-2)} - \frac{(\hat{\mu}^{(k-1)} - \hat{\mu}^{(k-2)})^2}{\hat{\mu}^{(k)} - 2\hat{\mu}^{(k-1)} + \hat{\mu}^{(k-2)}}$$

Inverse Power Method

For $q \in \mathbb{R}$, if λ_i are eigenvalues of A with eigenvectors of $(A-qI)^{-1}$.

Goal: find eigenvalues closest to q.

Let λ_J be the eigenvalues of A closest to q, then $|\lambda_J - q|^{-1}$ can be much larger than $|\lambda_j - q|^{-1}$ for $j \neq J$.

Algorithm 3 Inverse Power Method

- Choose q and $\vec{x}^{(0)} \neq 0$ s.t. $\|\vec{x}^{(0)}\| = 1$ and $a_1 \neq 0$
- Solve $(A - qI_n)\vec{w}^{(k)} = \vec{x}^{(k+1)}$ for $\vec{w}^{(k)}$

$$\vec{x}^{(k)} = \frac{\vec{w}^{(k)}}{\|\vec{w}^{(k)}\|} \quad \text{normalize}$$

$$\mu^{(k)} = \vec{x}^{(k)T} A \vec{x}^{(k)} \quad \text{Rayleigh Quotient}$$

If $|\mu^{(k)} - \mu^{(k-1)}| \leq$ tol. or $\|\vec{x}^{(k)} - \vec{x}^{(k-1)}\| <$ tol.,

then stop, return $\vec{x}^{(k)}$ and $\mu^{(k)}$.

eigenvalues eigenvectors

Rmk. Convergence of Inverse Power Method is linear but with more control on convergence by choosing q. The closer q is to an eigenvalue λ_k, the faster the convergence since the convergence rate is

$$O\left(\left|\frac{(\lambda-q)^{-1}}{(\lambda_j-q)}\right|^k\right) \quad \text{in general}$$

$$O\left(\left|\frac{(\lambda-q)}{(\lambda_j-q)}\right|^{2k}\right) \quad \ell_2\text{-norm, symmetric matrices}$$

where λ is the eigenvalue of A second closest to q in magnitude.

9.2 Householder's Method

Goal: convert symmetric matrices into tridiagonal matrices

Householder Transformations

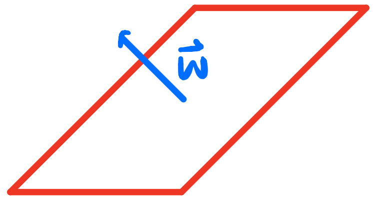
Def. Let $\vec{w} \in \mathbb{R}^n$ with $\|\vec{w}\|_2 = 1$. The matrix

$$P = I_n - 2\vec{w}\vec{w}^T$$

is called a **Householder matrix**. The linear transformation $\vec{x} \to P\vec{x}$ is a **Householder transformation**.

Rmk. 1. If $0 \neq \vec{w} \in \mathbb{R}^n$, then in general
$P = I_n - \frac{2\vec{w}\vec{w}^T}{\vec{w}^T\vec{w}}$ is a Householder matrix.

2. P defines a reflection about plane normal to $\vec{w}$



3. P depends on $\vec{w}$.

Thm. A Householder matrix $P = I_n - 2\vec{w}\vec{w}^T$ is symmetric and ON, which implies that $P^T = P = P^{-1}$.

Pf. $P^T = (I - 2\vec{w}\vec{w}^T)^T = I - 2\vec{w}\vec{w}^T = P$ (symmetry)

$$PP^T = (I - 2\vec{w}\vec{w}^T)(I - 2\vec{w}\vec{w}^T)^T$$
$$= I - 4\vec{w}\vec{w}^T + 4\vec{w}(\vec{w}^T\vec{w})\vec{w}^T$$
$$= I$$

Thus, we have $P^T = P = P^{-1}$. $\square$

Prop. Let $\vec{u} = \vec{x} + \|\vec{x}\|_2 \vec{e}_1$, where

$\vec{e}_1 = (1, 0, 0, \ldots, 0)^T \in \mathbb{R}^n$, and

$P = I_n - \frac{2\vec{u}\vec{u}^T}{\vec{u}^T\vec{u}}$. Then

$P\vec{x} = -\|\vec{x}\|_2 \vec{e}_1$

Pf. Note that

$$\begin{aligned}\vec{u}^T\vec{u} &= (\vec{x} + \|\vec{x}\|_2\vec{e}_1)^T(\vec{x} + \|\vec{x}\|_2\vec{e}_1) \\ &= \|\vec{x}\|_2^2 + 2\|\vec{x}\|_2\vec{e}_1^T\vec{x} + \|\vec{x}\|_2^2 \\ &= 2(\underbrace{\vec{x} + \|\vec{x}\|_2\vec{e}_1}_{\vec{u}})^T\vec{x} \\ &= 2\vec{u}^T\vec{x}\end{aligned}$$

$$\begin{aligned}P\vec{x} &= \vec{x} - \frac{2\vec{u}\vec{u}^T}{\vec{u}^T\vec{u}}\vec{x} \\ &= \vec{x} - \vec{u} = -\|\vec{x}\|_2\vec{e}_1 \quad \square\end{aligned}$$

Rmk. 1. This shows that the Householder matrix (with choice $\vec{u} = \vec{x} + \|\vec{x}\|_2\vec{e}_1$) zeroes all components of a given $\vec{x}$ except for first one.

2. Note $\vec{x} + \|\vec{x}\|_2\vec{e}_1 = \begin{bmatrix} x_1 + \|\vec{x}\|_2 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

if $x_1 < 0$ and $x_1 \approx \|\vec{x}\|_2$, we can get significant cancellation...

To avoid subtraction of 2 nearly

equal numbers in the first component of $\vec{x}+\|\vec{x}\|_2\vec{e}_1$, we choose $\vec{u}=\vec{x}+\operatorname{sgn}(x_1)\cdot\|\vec{x}\|_2\vec{e}_1$ for stability.

3. Since $\vec{x}^T P^T = -\|\vec{x}\|_2 \vec{e}_1^T$, we can extend results to row vectors.

Householder's Method

Def. A matrix $H\in\mathbb{R}^{n\times n}$ is called

- upper Hessenberg Matrix if $h_{ij}=0$ for all $i\geq j+2$

$$\begin{bmatrix} x & x & x & \cdots & x \\ x & x & x & & \vdots \\ 0 & x & x & \ddots & \\ 0 & 0 & x & & \\ \vdots & \vdots & 0 & & x \\ 0 & 0 & 0 & \cdots 0 \; x & x \end{bmatrix}$$

- lower Hessenberg Matrix if $h_{ij}=0$ for all $j\geq i+2$

$$\begin{bmatrix} x & x & 0 & 0 & \cdots & 0 \\ x & x & x & 0 & & \vdots \\ x & x & x & x & & \\ \vdots & \vdots & & \ddots & \ddots & 0 \\ & & & & & x \\ x & x & x & \cdots & x & x \end{bmatrix}$$

Rmk. A tridiagonal matrix is upper and lower Hessenberg. Assume that $A\in\mathbb{R}^{n\times n}$ is symmetric. We look for a tridiagonal matrix T

such that $A = \underbrace{Q^T T Q}_{\text{similarity transformation } (Q^T = Q^{-1})}$.

Let $\vec{x}^{(1)} = (a_{21}, \ldots, a_{n1})^T$

$$\vec{v}^{(1)} = \begin{cases} \vec{x}^{(1)} + \operatorname{sgn}(x_1^{(1)}) \cdot \|\vec{x}\|_2 \vec{e}_1 & \text{if } x_1^{(1)} \neq 0 \\ \vec{x}^{(1)} + \|\vec{x}\|_2 \vec{e}_1 & \text{if } x_1^{(1)} = 0 \end{cases}$$

$$P_1 = I_{n-1} - 2\frac{\vec{v}^{(1)}\vec{v}^{(1)T}}{\|\vec{v}^{(1)}\|_2^2}$$

Then $\underset{(n-1)\times(n-1)}{P_1} \underset{(n-1)\times n}{A(2:n, 1:n)} = \underset{(n-1)\times n}{\begin{bmatrix} * & * & \cdots & * \\ 0 & * & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & * & * & * \end{bmatrix}}$

where $A(i:i', j:j')$ denotes $(i'-i+1)\times(j'-j+1)$ submatrix of A.

Likewise, we can get $P_k \in \mathbb{R}^{(n-k)\times(n-k)}$ for $k = 1, \ldots, n-2$ and define

$$Q_k = \begin{bmatrix} I_k & \\ & P_k \end{bmatrix} \in \mathbb{R}^{n\times n}$$

and $Q = Q_{n-2} Q_{n-3} \ldots Q_2 Q_1$

which is orthogonal

Then QA is upper Hessenberg, and QAQ^T is a tridiagonal matrix as desired.

Pictorally:

$$A = \begin{bmatrix} * & x & x & x \\ * & x & x & x \\ * & x & x & x \\ * & x & x & x \end{bmatrix} \xrightarrow{Q_1 A Q_1^T} \begin{bmatrix} x & x & 0 & 0 \\ x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}$$

$$\xrightarrow{Q_2 Q_1 A Q_1^T Q_2^T} \begin{bmatrix} x & x & 0 & 0 \\ x & x & x & 0 \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix}$$ tridiagona l

Alg. 9.5 Householder Reduction to Tridiagonal (Hessenberg) Matrix

- Initialize $T = A$, $Q = I$.
- For $k = 1, 2, \dots n-2$:

$\vec{x} = T(k+1:n, k) \leftarrow k^{th}$ col

$$\vec{v}^{(k)} = \begin{cases} \vec{x} + \mathrm{sgn}(x_1) \cdot \|\vec{x}\|_2 \vec{e}_1 & \text{if } x_1 \neq 0 \\ \vec{x} + \|\vec{x}\|_2 \vec{e}_1 & \text{if } x_1 = 0 \end{cases}$$

$$\vec{v}^{(k)} = \frac{\vec{v}^{(k)}}{\|\vec{v}^{(k)}\|_2}$$

$$P_k = I_{n-k} - 2\vec{v}^{(k)} \vec{v}^{(k)T} \quad \in \mathbb{R}^{(n-k)\times(n-k)}$$

$$Q_k = \begin{bmatrix} I_k & \\ & P_k \end{bmatrix}$$

left HH transform

$$T(k+1:n, k:n) \leftarrow T(k+1:n, k:n) - 2\vec{v}^{(k)}(\vec{v}^{(k)T} T(k+1:n, k:n))$$

$$T(k:n, k+1:n) \leftarrow T(k:n, k+1:n) - 2(\vec{v}^{(k)T} T(k:n, k+1:n))\vec{v}^{(k)T}$$

right HH transform

$(\Leftrightarrow T \leftarrow Q_k T Q_k^T)$

$$Q \leftarrow Q_k Q$$

Rmk. If A is not symmetric, then Alg. 9.5 without right HH transform returns an upper Hessenberg matrix.