Here's a summary of Power method. QR method and Simultaneous method.

Power method.

Assume the matrix $A \in \mathbb{C}^{n\times n}$ is diagonalizable (i.e. $A = QDQ^{-1}$. where D is a diagonal matrix, not a general Jordan block). we denote the eigenpairs by $(\lambda_i, \vec{q_i})$. $A\vec{q_i} = \lambda_i \vec{q_i}$ and its eigenvalues can be ordered in such a way that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$

QR method is defined via.

given $\vec{x}^{(0)} = \sum_{i=1}^{n} a_i \cdot \vec{q_i}$, compute $\vec{x}^{(k+1)} = \dfrac{A\vec{x}^{(k)}}{\|A\vec{x}^{(k)}\|_\infty}$ as long as

$a_1 \neq 0$. $\vec{x}^{(k)} \to \dfrac{\vec{q_0}}{\|\vec{q_0}\|_\infty}$

Comment: ① $a_1 \neq 0$. and $|\lambda_1| > |\lambda_2|$ are essential for the convergence of $\vec{x}^{(k)}$.

② in each iteration, "normalizing" ($A\vec{x}^{(k)} \to \dfrac{A\vec{x}^{(k)}}{\|A\vec{x}^{(k)}\|_\infty}$) is for numerical stability. we can change it to any other norm. like $\|A\vec{x}^{(k)}\|_2$.

No matter what kind of normalization factor is chosen,

$\vec{x}^{(k)} = C_k A^k \vec{x}^{(0)}$. $\vec{x}^{(k)}$ is a scaled up version of $A^k \vec{x}^{(0)}$.

As for the scaling factor. C_k. here's a trick to determine it.

e.g. in the case $\|\cdot\|_\infty$. we know by definition. $\|\vec{x}^{(k)}\|_\infty = 1$.

so $C_k = \dfrac{1}{\|A^k \vec{x}^{(0)}\|_\infty}$.

In general. if $\vec{x}^{(k)} = \dfrac{A\vec{x}^{(k-1)}}{\|A\vec{x}^{(k-1)}\|_*}$ for some $\|\cdot\|_*$, then.

$$C_k = \|A^k \vec{x}^{(0)}\|_*.$$

For the inverse iteration method. our assumption on A is.

1. A is diagonalizable.
2. A is Invertible
3. eigenvalues of A are $\frac{1}{\lambda_n}, \cdots, \frac{1}{\lambda_1}$, they satisfy

$$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \geq \cdots \geq \left|\frac{1}{\lambda_1}\right|$$

Let $\vec{x}^{(0)} = \sum_1^n a_i \cdot \vec{q_i}$. $\vec{x}^{(k)} = \frac{A^{-1}x^{(k-1)}}{\|A^{-1}x^{(k-1)}\|_\infty}$, as long as $a_n \neq 0$. then

$$\vec{x}^{(k)} \to \frac{\vec{q_n}}{\|\vec{q_n}\|_\infty}$$

For the inverse power method with shift. its goal is to find the eigenvalue of A that's closest to μ. it's to do inverse power method to $A - \mu I$.

Convergence Rate of above three methods. is both.

$$\frac{\text{something second largest}}{\text{something the largest}}$$

"something" ~ magnitude of the eigenvalues of iteration matrix.

Power method: iteration matrix is A. so its largest and second largest. magnitude eigenvalues are $|\lambda_1|$, $|\lambda_2|$. rate $= \frac{|\lambda_2|}{|\lambda_1|}$

Inverse Power method:

iteration matrix is A^{-1}. so its largest and second largest. magnitude eigenvalues are $|\lambda_n|^{-1}$. $|\lambda_{n-1}|^{-1}$. rate $= \frac{|\lambda_{n-1}|^{-1}}{|\lambda_n|^{-1}} = \left|\frac{\lambda_n}{\lambda_{n-1}}\right|$

Inverse Power method with shift:

iteration matrix is $(A-\mu I)^{-1}$ so its largest and second largest magnitude eigenvalues are some $|\lambda_i-\mu|^{-1}$, $|\lambda_j-\mu|^{-1}$ for some i, j.

rate is $\dfrac{|\lambda_j-\mu|^{-1}}{|\lambda_i-\mu|^{-1}}$

There's a type of question related to Inverse Power method with shift

that's, telling you information of A. ask you to design μ s.t. the method converges the fastest.

In this case your task is to minimize $\dfrac{|\lambda_j-\mu|^{-1}}{|\lambda_i-\mu|^{-1}}$

A more advanced method than Inverse Power method with shift is Rayleigh Quotient.

its intuition is to adjust the shift μ according to the latest $\vec{x}^{(k)}$.

It can achieve qudratic convergence rate, but its analysis is much more complicated. which is beyond the scope of this course.

Next. we discuss simultaneous method and QR method.

In this section. we assume.

1. A is symmetric. real.
2. A is nonsingular.

These two assumptions is to reduce our case to the easiest one.

Because we know. for nonsingular A. we have QR decomposition.

$A=QR$, with Q containing orthonormal columns, R is upper triangular with positive diagonal elements.

$$A = \left(\vec{q_1} \,|\, \vec{q_2} \,|\, \cdots \,|\, \vec{q_n} \right) \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ \underbrace{0 \quad 0}_{} & & r_{33} & \\ & & & r_{nn} \end{pmatrix}$$

$Q^TQ = I_n$, $\vec{q_i}^T \vec{q_j} = \delta_{ij}$. $r_{ii} > 0$, $\forall i = 1, 2, \cdots, n$.

Another advantage is for a real symmetric matrix. its eigendecomposition is

$A = \tilde{Q} D \tilde{Q}^T$. where $\tilde{Q} = \left(\vec{\tilde{q}_1} \,|\, \cdots \,|\, \vec{\tilde{q}_n} \right)$

$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ $\lambda_j \in \mathbb{R}$.

and $\vec{\tilde{q}_i}^T \vec{\tilde{q}_j} = \delta_{ij}$. $\tilde{Q}^T \tilde{Q} = I$.

For geneal diagonalizable matrix. we only have $A = PDP^{-1}$.

its eigenvector is not necessarily orthogonal.

and eigenvalue is not necessarily real-valued.

But for real symmetric matrix. its eigenvector is orthogonal

and eigenvalue is real-valued.

However. Q obtained by QR decomposition and $\tilde{Q}$ obtained by eigendecomposition are NOT necessarily the same.

Above are some good properties brought by the assumptions.

First, we introduce simultaneous method.

Suppose $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \cdots$, $\vec{x}_1^{(0)}, \vec{x}_2^{(0)}$ are two initial vectors

we do power method to $\vec{v_1}, \vec{v_2}$ simultaneously.

$$\left(\vec{x}_1^{(k)}, \vec{x}_2^{(k)} \right) \propto \left(A^k \vec{x}_1^{(0)} \,|\, A^k \vec{x}_2^{(0)} \right)$$

Since A is assumed to be invertible, then $\left(\vec{x}_1^{(0)}, \vec{x}_2^{(0)} \right)$ are independent

$\Leftrightarrow \left(\vec{x}_1^{(k)}, \vec{x}_2^{(k)} \right)$ are independent.

but. $\vec{x}_1^{(0)} = a_1 \cdot \vec{q}_1 + \sum_2^n a_j \cdot \vec{q}_j$. $\vec{x}_2^{(0)} = b_1 \cdot \vec{q}_1 + \sum_2^n b_j \cdot \vec{q}_j$

as long as a_1, b_1 are both nonzero, then

limits of $\vec{x}_1^{(k)}, \vec{x}_2^{(k)}$ would be the same $\frac{\vec{q}_1}{\|\vec{q}_1\|_\infty}$ though.

$\vec{x}_1^{(k)}, \vec{x}_2^{(k)}$ are always independent for all k.

An example is $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix}$. $\vec{x}_1^{(0)} = (1,1,1)^T$, $\vec{x}_2^{(0)} = (1,2,3)^T$.

then $A^k \vec{x}_1^{(0)} = (1, 2^{-k}, 4^{-k})^T$ $A^k \vec{x}_2^{(0)} = (1, 2^{-k+1}, 3 \cdot 4^{-k})^T$

$$\vec{x}_1^{(k)} = \frac{A^k \vec{x}_1^{(0)}}{\|A^k \vec{x}_1^{(0)}\|_\infty} = (1, 2^{-k}, 4^{-k})^T \to (1,0,0)$$

$$\vec{x}_2^{(k)} = \frac{A^k \vec{x}_2^{(0)}}{\|A^k \vec{x}_2^{(0)}\|_\infty} = (1, 2^{-k+1}, 3 \cdot 4^{-k})^T \to (1,0,0)$$

Similar observation occurs for the case $\lambda_1 = \lambda_2 = \cdots = \lambda_k > |\lambda_{k+1}| \geq \cdots \geq |\lambda_n|$

let $eig(\lambda_1) = span\{\vec{v} \mid A\vec{v} = \lambda_1 \vec{v}\}$. then each vector $\vec{w}$ can be written as $\vec{w} = \vec{v} + \vec{u}$. where $\vec{v} \in eig(\lambda_1)$, $\vec{u} \in eig(\lambda_1)^\perp$

$eig(\lambda_1)^\perp = span\{\vec{w} \mid \vec{w}^T \vec{v} = 0 \ \forall \vec{v} \in eig(\lambda_1)\}$.

As long as the initial $\vec{x}^{(0)} = \vec{v}^{(0)} + \vec{u}^{(0)}$ satisfies $\vec{v}^{(0)} \neq 0$. then.

its iteration result $\vec{x}^{(k)} = \frac{A^k \vec{x}^{(0)}}{\|A^k \vec{x}^{(0)}\|_\infty} \to$ some element in $eig(\lambda_1)$

Even if you choose $\vec{x}_1^{(0)}, \cdots, \vec{x}_k^{(0)}$ k independent initial vectors,

and decompose $\vec{x}_j^{(0)} = \vec{v}_j^{(0)} + \vec{u}_j^{(0)}$. each $\vec{v}_j^{(0)} \neq 0$.

Then $(\vec{x}_1^{(m)}, \cdots, \vec{x}_k^{(m)}) \propto (A^m \vec{x}_1^{(0)}, \cdots, A^m \vec{x}_k^{(0)})$ is always independent.

and each $\vec{x}_j^{(m)} \to$ some element in $eig(\lambda_1)$ as $m \to \infty$.

However, we cannot guarantee that

$\left(\lim_{m\to\infty}\vec{x}_1^{(m)}, \lim_{m\to\infty}\vec{x}_2^{(m)}, \cdots, \lim_{m\to\infty}\vec{x}_k^{(m)}\right)$ must be independent.

and form a basis of $eig(\lambda_1)$.

A strategy to solve this problem when doing power methods to multiple initial vectors simultaneously is that.

$$X^{(0)}=(\vec{x}_1^{(0)}|\cdots|\vec{x}_k^{(0)}) \to X^{(1)}=AX^{(0)} \to X^{(2)}=A^2X^{(0)} \to \cdots$$

add one more step

$X^{(1)}=$ QR-factorization of $X^{(1)}$

Because by above analysis, columns of $X^{(k)}$ being always independent is not enough to make their limit vectors distinct. That's why we add one step to make $X^{(k)}$ orthogonal in each step.

Algorithm.

$X^{(0)}=(\vec{x}_1^{(0)}|\cdots|\vec{x}_n^{(0)})$. Do QR factorization of $X^{(0)}$. $X^{(0)}=\bar{Q}^{(0)}R^{(0)}$

Start with $\bar{Q}^{(0)}$, For $k=1,2,\cdots$ do the following steps recursively until reaching some stopping criterion.

1. $W=A\bar{Q}^{(k-1)}$ 2. QR factorization: $W=\bar{Q}^{(k)}R^{(k)}$

Some observations:

1. $\bar{Q}^{(k)}$ is the "Q" matrix obtained by doing QR factorization to $A^kX^{(0)}$

This can be prove by induction.

It's clearly true for $k=0$. Now assume $A^k\bar{Q}^{(0)}=\bar{Q}^{(k)}\bar{R}^{(k)}$.

if we let $w=A\bar{Q}^{(k)}=\bar{Q}^{(k+1)}R^{(k+1)}$. (algorithm did).

$$A^{k+1}\overline{Q}^{(0)} \underset{\substack{\uparrow \\ \text{assumption}}}{=} A\overline{Q}^{(k)}\overline{R}^{(k)} = W\cdot\overline{R}^{(k)} = \overline{Q}^{(k+1)}R^{(k+1)}\overline{R}^{(k)}$$

So let $\overline{R}^{(k+1)} \stackrel{!}{=} R^{(k+1)}\overline{R}^{(k)}$, then $(\overline{Q}^{(k+1)}, \overline{R}^{(k+1)})$ is QR factorization of $A^{k+1}\overline{Q}^{(0)}$

So. $(\overline{Q}^{(1)}, R^{(1)}), (\overline{Q}^{(2)}, R^{(2)}), \cdots, (\overline{Q}^{(k)}, R^{(k)})$ are intermediate quantities obtained in the algorithm, then.

$\overline{R}^{(k)} = R^{(k)}\cdot R^{(k-1)}\cdots R^{(0)}$, and $\overline{Q}^{(k)}$ is the QR factorization of $A^k X^{(0)}$

How to understand it?

$X^{(0)} = \overline{Q}^{(0)}R^{(0)}$. let $X^{(0)} = (\vec{x}_1^{(0)}|\cdots|\vec{x}_n^{(0)})$ then.

$\overline{Q}^{(0)} = (\vec{q}_1^{(0)}|\cdots|\vec{q}_n^{(0)})$ is an orthonormal basis for span$\{\vec{x}_j^{(0)} | j=1\cdots n\}$

And $A\overline{Q}^{(0)}$ is a basis for span$\{A\vec{x}_j^{(0)} | j=1,\cdots,n\}$.

Note that $A\overline{Q}^{(0)}$'s columns are not orthonormal.

However, by doing QR factorization to $A\overline{Q}^{(0)}$,

$\overline{Q}^{(1)}$'s columns are an orthonormal basis of span$\{A\vec{x}_j^{(0)} | j=1,\cdots,n\}$.

Repeating this process.

$\overline{Q}^{(k)}$'s columns are an orthonormal basis of span$\{A^k\vec{x}_j^{(0)} | j=1,\cdots,n\}$.

That implies that $\overline{Q}^{(k)}$ might be the "Q" matrix obtained in QR factorization of $A^k X^{(0)}$

And it turns out that this is true

2. By some theory, assume

① A is symmetric.

② $|\lambda_1| \geq \cdots \geq |\lambda_k| > |\lambda_{k+1}| \geq \cdots \geq |\lambda_n|$

③ let P be the orthogonal projection onto the span of the first k eigenvectors, and $\bar{Q}^{(0)} = (\vec{q}_1^{(0)} | \cdots | \vec{q}_k^{(0)} | \cdots)$.

Assume $P(\vec{q}_1^{(0)} | \cdots | \vec{q}_k^{(0)})$ is injective, namely projection of first k column vectors in $\bar{Q}^{(0)}$ is independent.

Then we have

$$\| A Q_k^{(m)} - Q_k^{(m)} \Lambda_{kk}^{(m)} \|_2 \leq C_0 \cdot \left(\frac{|\lambda_{k+1}|}{|\lambda_k|} \right)^m,$$

$$Q_k^{(m)} = (\vec{q}_1^{(m)} | \cdots | \vec{q}_k^{(m)}) \text{ and } \Lambda_{kk}^{(m)} = \left(Q_k^{(m)}\right)^T A Q_k^{(m)}$$

Very abstract statement? Yes, but we can try to interpret it intuitively.

let $(\lambda_j, \vec{q}_j)$ be all eigen-pairs of A. then $\left(A\vec{q}_j = \lambda_j \cdot \vec{q}_j\right)_{j=1}^{k}$ can be rewritten as

$$A(\vec{q}_1 | \cdots | \vec{q}_k) = (\vec{q}_1 | \cdots | \vec{q}_k) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{pmatrix} \quad \cdots (\star)$$

since A is symmetric, $q_i^T q_j = \delta_{ij}$. which means $\vec{q}_i^T A \vec{q}_j = \delta_{ij} \cdot \lambda_j$.

equation $(\star) \Leftrightarrow A Q_k = Q_k \cdot (Q_k^T A Q_k) \qquad Q_k = (\vec{q}_1 | \cdots | \vec{q}_k)$.

Very similar to $A Q_k^{(m)} - Q_k^{(m)} ({Q_k^{(m)}}^T A Q_k^{(m)})$, right?

Yes. this theory is telling us, under some assumptions,

$Q_k^{(m)}$ would converge to $(\vec{q}_1 | \cdots | \vec{q}_k)$.

${Q_k^{(m)}}^T A Q_k^{(m)}$ would converge to $\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{pmatrix}$.

So what simultaneous iteration did is doing QR factorization on $A^k X^{(0)}$.

and theory guarantee that $Q_k^{(m)}$ would converge to a diagonal matrix. $\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_k\end{pmatrix}$.

Usually we choose $X^{(0)} = I_n$. identity matrix, and we have $A^k = \bar{Q}^{(k)} \bar{R}^{(k)}$

So what's its connection with QR method?

By above. we know $A^{(m)} = \bar{Q}^{(m)^T} A \bar{Q}^{(m)}$. is also of interest. because hopefully $A^{(m)}$ would converge to $\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$

Sometimes we are more interested in computing the eigenvalues (λ_j) than computing the eigen-vectors (q_j).

So is there any direct way to compute $A^{(m)}$. so that we can directly obtain $\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$?

Answer is Yes! By QR method.

Next. we start from simultaneous method. which outputs $(\bar{Q}^{(k)}, R^{(k)})_{k=0}^{\infty}$.

try to derive a more efficient way to compute $A^{(k)} = \bar{Q}^{(k)^T} A \bar{Q}^{(k)}$

and finally the QR method.

Remember we have $\bar{Q}^{(m+1)} R^{(m+1)} = A \bar{Q}^{(m)}$.

Left multiply both sides by $(\bar{Q}^{(m)})^T$.

$$\underbrace{\left(\bar{Q}^{(m)^T} \bar{Q}^{(m+1)}\right) R^{(m+1)}}_{} = A^{(m)} = \bar{Q}^{(m)^T} A \bar{Q}^{(m)}$$

Surprisingly, it's QR factorization of $A^{(m)}$!

So if we denote $Q_{QR}^{(k+1)} = \bar{Q}^{(k)T} \bar{Q}^{(k+1)}$, $R_{QR}^{(k+1)} = R^{(k+1)}$.

we have $Q_{QR}^{(m+1)} R_{QR}^{(m+1)} = A^{(m)}$

(To avoid abuse of notation. here we clarify again, $\bar{Q}^{(k)}$. $R^{(k)}$ are quantites obtained in simultaneous method, $\bar{R}^{(k)} = R^{(k)} \cdots R^{(0)}$. $A^k = \bar{Q}^{(k)} \bar{R}^{(k)}$. and $A^{(k)} = \bar{Q}^{(k)T} A \bar{Q}^{(k)}$)

Define $Q_{QR}^{(k+1)}$ in this way, then $\bar{Q}^{(k)} \cdot Q_{QR}^{(k+1)} = \bar{Q}^{(k+1)}$

$$
\begin{aligned}
A^{(k+1)} = \bar{Q}^{(k+1)T} A \bar{Q}^{(k+1)} &= Q_{QR}^{(k+1)T} \underbrace{\bar{Q}^{(k)T} A \bar{Q}^{(k)}} Q_{QR}^{(k+1)} \\
&= Q_{QR}^{(k+1)T} A^{(k)} Q_{QR}^{(k+1)} \\
&= Q_{QR}^{(k+1)T} Q_{QR}^{(k+1)} R_{QR}^{(k+1)} Q_{QR}^{(k+1)} \\
&= R_{QR}^{(k+1)} Q_{QR}^{(k+1)}
\end{aligned}
$$

In summary. $A^{(m)} = Q_{QR}^{(m+1)} R_{QR}^{(m+1)}$

$A^{(m+1)} = R_{QR}^{(m+1)} Q_{QR}^{(m+1)}$

QR method: Let $A_{QR}^{(0)} = A$.

For $k = 1, 2 \cdots$ obtain QR factorization of

$A_{QR}^{(k-1)} = Q_{QR}^{(k)} R_{QR}^{(k)}$

Then let $A_{QR}^{(k)} = R_{QR}^{(k)} Q_{QR}^{(k)}$. repeat the above step.

So by above, we have. $A_{QR}^{(k)} = A^{(k)} = \bar{Q}^{(k)T} A \bar{Q}^{(k)}$

And $Q_{QR}^{(k+1)} = \bar{Q}^{(k)T} \bar{Q}^{(k+1)}$, $R_{QR}^{(k+1)} = R^{(k)}$

$$\Rightarrow \bar{Q}_{QR}^{(k)} := Q_{QR}^{(1)} \cdots Q_{QR}^{(k)}$$

$$= \underbrace{\bar{Q}^{(0)T}}_{\substack{\vdots \\ I \text{ if we let } X^{(0)} = I.}} \underbrace{\bar{Q}^{(1)} \bar{Q}^{(1)T}}_{I} \underbrace{\bar{Q}^{(2)} \cdots}_{I} \underbrace{\bar{Q}^{(k-1)} \bar{Q}^{(k-1)T}}_{I} \bar{Q}^{(k)}$$

$$= \bar{Q}^{(k)}$$

also if we choose $X^{(0)} = I_n$ in the simultaneous method. then $\bar{Q}^{(0)}, \bar{R}^{(0)} = I_n$.

$$A^k = \bar{Q}^{(k)} \bar{R}^{(k)} = \bar{Q}_{QR}^{(k)} \cdot (R^{(k)} \cdots R^{(1)} R^{(0)})$$

$$= \bar{Q}_{QR}^{(k)} \underbrace{(R_{QR}^{(k)} \cdots R_{QR}^{(1)})}_{\triangleq \bar{R}_{QR}^{(k)}}$$

$$= \bar{Q}_{QR}^{(k)} \bar{R}_{QR}^{(k)}$$

In conclusion. the workflow is

Do Power method to multiple vectors simultaneously

$\Rightarrow$ propose Simultaneous Method. with output $(\bar{Q}^{(k)}, R^{(k)})$. $X^{(0)} = I_n$.

$\Rightarrow$ important quantities

$$\bar{R}^{(k)} := R^{(k)} R^{(k-1)} \cdots R^{(1)}, \quad A^{(k)} := \bar{Q}^{(k)T} A \bar{Q}^{(k)}$$

Important equations.

$$A^k = \bar{Q}^{(k)} \bar{R}^{(k)}$$

$\Rightarrow$ compute $A^{(k)}$ more efficiently thus propose QR method. with outputs.

$(Q_{QR}^{(k)}, R_{QR}^{(k)}, A^{(k)})$

$\Rightarrow$ important quantities.

$$\overline{Q}_{QR}^{(k)} := Q_{QR}^{(1)} \cdots Q_{QR}^{(k)}, \quad \overline{R}_{QR}^{(k)} = R_{QR}^{(k)} \cdots R_{QR}^{(1)}$$

important equations:

$$\overline{Q}_{QR}^{(k)} = \overline{Q}^{(k)}, \quad \overline{R}_{QR}^{(k)} = \overline{R}^{(k)}$$

$$A^{(k)} = \overline{Q}_{QR}^{(k)T} A \, \overline{Q}_{QR}^{(k)}.$$

$$A^{k} = \overline{Q}_{QR}^{(k)} \overline{R}_{QR}^{(k)} = \overline{Q}^{(k)} \overline{R}^{(k)}$$

Last thing is QR / Simultaneous method with shift. as well as the connection between QR method and Inverse power method.

Remember $A^k = \overline{Q}^{(k)} \overline{R}^{(k)}$. so $(A^{-1})^k = (\overline{R}^{(k)})^{-1} (\overline{Q}^{(k)})^{-1}$

$$= \underline{\underline{(\overline{R}^{(k)})^{-1}}} \, \overline{Q}^{(k)T}$$

inverse of upper triangular matrix is still upper triangular.

Since A is symmetric. $(A^{-1})^k = \left((A^{-1})^k\right)^T = \overline{Q}^{(k)} \underbrace{\left((\overline{R}^{(k)})^{-1}\right)^T}_{\text{lower triangular}}$

Then right multiply both sides by $P = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \iddots & \vdots \\ 0 & 1 & 1 & & \vdots \\ 1 & 0 & 0 & & 0 \end{pmatrix} = (\vec{e}_n | \vec{e}_{n-1} | \cdots | \vec{e}_1)$

$$A^{-k} P = [\overline{Q}^{(k)} P] \, [P (\overline{R}^{(k)})^{-T} P]$$

$$= [\vec{q}_n | \cdots | \vec{q}_1] \, \underline{[P (\overline{R}^{(k)})^{-T} P]}.$$

$(\overline{R}^{(k)})^{-T}$: Δ $\rightarrow P (\overline{R}^{(k)})^{-T}$: ∇

lower

$\rightarrow P (\overline{R}^{(k)})^{-T} P$: ∇ upper

This is a QR factorization of $A^{-k}P$.

So simultaneous iteration applied to $X^{(0)} = I$. $\Leftrightarrow$.

simultaneous inverse iteration to "flipped" I. namely P

because $A^{-k}P$. $A^k I$ obtain same $\bar{Q}^{(k)}$. just in different column order.

Another more common statement is

The last column of simultaneous method $\left(A^k \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \propto \vec{q_n} \right)$.

is equivalent to inverse method to $\left(A^{-k} \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \propto \vec{q_n} \right)$

Since QR algorithm behaves like inverse iteration. it's natural to introduce shift $\mu^{(k)}$ to accerlerate the convergence.

$$\begin{cases} A^{(k-1)} - \mu^{(k)} I = Q_{QR}^{(k)} R_{QR}^{(k)} \\ A^{(k)} = R_{QR}^{(k)} Q_{QR}^{(k)} + \mu^{(k)} I. \end{cases}$$

How to select $\mu^{(k)}$? remember last column of QR method is doing inverse iteration, so $\mu^{(k)}$ should be close to λ_n in order to accerlerate convergence. that's why

$$\mu^{(k)} := \frac{(\vec{q}_n^{(k)})^T A (\vec{q}_n^{(k)})}{(\vec{q}_n^{(k)})^T (\vec{q}_n^{(k)})}$$

OK. Above is an illustration of QR, simultaneous, power method.

Hopefully, it would be easier for you to prove the several properties listed in the lecture.

if following the workflow in this note.

More clearly. ① prove $A^k = \bar{Q}^{(k)} \bar{R}^{(k)}$ for simultaneous method.

② introduce $A^{(k)} = \bar{Q}^{(k)T} A \bar{Q}^{(k)}$.

③ define $Q_{QR}^{(k+1)} = \bar{Q}^{(k)T} Q^{(k+1)}$, $R_{QR}^{(k)} = R^{(k)}$.

And prove that $A^{(m)} = Q_{QR}^{(m+1)} R_{QR}^{(m+1)}$

$A^{(m+1)} = R_{QR}^{(m+1)} Q_{QR}^{(m+1)}$

④ show that $Q_{QR}^{(k)}$, $R_{QR}^{(k)}$ defined in this way coincide with QR method.

⑤ Define $\bar{Q}_{QR}^{(k)} := Q_{QR}^{(1)} \cdots Q_{QR}^{(k)}$. $\bar{R}_{QR}^{(k)} = R_{QR}^{(k)} \cdots R_{QR}^{(1)}$

we can prove. finally

$\bar{Q}_{QR}^{(k)} = \bar{Q}^{(k)}$, $\bar{R}_{QR}^{(k)} = \bar{R}^{(k)}$

$A^{(k)} = \bar{Q}_{QR}^{(k)T} A \bar{Q}_{QR}^{(k)}$.

$A^k = \bar{Q}_{QR}^{(k)} \bar{R}_{QR}^{(k)} = \bar{Q}^{(k)} \bar{R}^{(k)}$

Of course, in the exam, you usually don't have time to really follow all above steps.

What if you are just directly given QR, simultaneous method, and asked to prove properties listed in the course note?

My suggestion is that. ①. prove $A_{QR}^{(k)} = Q_{QR}^{(k)T} A_{QR}^{(k-1)} Q_{QR}^{(k)}$.

thus $A_{QR}^{(k)} = \bar{Q}_{QR}^{(k)T} A \bar{Q}_{QR}^{(k)}$

(proof of these two equations even don't need induction.
just rely on. QR method)

② . show
$$\begin{cases} A^k = \bar{Q}^{(k+1)} \bar{R}^{(k+1)} \\ A^k = \bar{Q}_{QR}^{(k+1)} \bar{R}_{QR}^{(k+1)}. \end{cases}$$

To prove these two, you need induction. i.e.

Assume $\begin{cases} A^{k-1} = \bar{Q}^{(k)} \bar{R}^{(k)} \\ A^{k-1} = \bar{Q}_{QR}^{(k)} \bar{R}_{QR}^{(k)} \end{cases}$, show $\begin{cases} A^k = \bar{Q}^{(k+1)} \bar{R}^{(k+1)} \\ A^k = \bar{Q}_{QR}^{(k+1)} \bar{R}_{QR}^{(k+1)} \end{cases}$

For the proof of

$$A^{k-1} = \bar{Q}^{(k)} \bar{R}^{(k)} \Rightarrow A^k = \bar{Q}^{(k+1)} \bar{R}^{(k+1)}$$

it's easy, clear. so I skip it here.

I think more tedious pare is, to prove

$$A^{k-1} = \bar{Q}_{QR}^{(k)} \bar{R}_{QR}^{(k)} \Rightarrow A^k = \bar{Q}_{QR}^{(k+1)} \bar{R}_{QR}^{(k+1)}$$

Here I provide another proof for you, a bit different from the process showed in the lecture.

First step is always $A^k = A \cdot A^{k-1} = A \, \bar{Q}_{QR}^{(k)} \bar{R}_{QR}^{(k)}$.

Secondly. remember. we have. $A_{QR}^{(k+1)} = \bar{Q}_{QR}^{(k+1)T} A \, \bar{Q}_{QR}^{(k+1)}$

So $A = \bar{Q}_{QR}^{(k+1)} A_{QR}^{(k+1)} \left(\bar{Q}_{QR}^{(k+1)}\right)^T$ (that's how $\bar{Q}_{QR}^{(k+1)}$ in the " $A^k = \bar{Q}_{QR}^{(k+1)} \bar{R}_{QR}^{(k+1)}$ was introduced)

Then $A^k = \underline{A}\, \overline{Q}_{QR}^{(k)} \overline{R}_{QR}^{(k)}$

$$= \underline{\overline{Q}_{QR}^{(k+1)} A_{QR}^{(k+1)} \left(\overline{Q}_{QR}^{(k+1)}\right)^T} \; \overline{Q}_{QR}^{(k)} \overline{R}_{QR}^{(k)}$$

Remember $\overline{Q}_{QR}^{(k+1)} = Q_{QR}^{(1)} \cdots Q_{QR}^{(k)} Q_{QR}^{(k+1)}$.

and $A_{QR}^{(k+1)} = \underline{R_{QR}^{(k+1)} Q_{QR}^{(k+1)}}$

So $$= \overline{Q}_{QR}^{(k+1)} A_{QR}^{(k+1)} \underbrace{\left(\overline{Q}_{QR}^{(k+1)}\right)^T \overline{Q}_{QR}^{(k)}}_{= Q_{QR}^{(k+1)\,T}} \overline{R}_{QR}^{(k)}$$

$$= \overline{Q}_{QR}^{(k+1)} R_{QR}^{(k+1)} \overline{R}_{QR}^{(k)} = \overline{Q}_{QR}^{(k+1)} \overline{R}_{QR}^{(k+1)}$$

3 so $\overline{Q}_{QR}^{(k)} = \overline{Q}^{(k)}$. $\overline{R}^{(k)} = \overline{R}_{QR}^{(k)}$.

and $A_{QR}^{(k)} = \overline{Q}_{QR}^{(k)\,T} A\, \overline{Q}_{QR}^{(k)} = \overline{Q}^{(k)\,T} A\, \overline{Q}^{(k)} = A^{(k)}$.