The QR factorization [We assume symm real matrices in following]

One other important matrix factorization that can be easily obtained with Householder reflectors is the QR facto:

let $A \in \mathbb{R}^{m\times n}$, then $\exists$ Q unitary $(Q^*Q = I)$ and R (∇) s.t.

$$A = QR$$

$$\underset{m\times n}{A} = \underset{m\times n}{Q}\;\underset{n\times n}{R}$$

Idea: apply Householder reflectors on the left of A to obtain upper triangular matrix:

$$\underset{A}{\begin{bmatrix} \times&\times&\times\\ \times&\times&\times\\ \times&\times&\times\\ \times&\times&\times\\ \times&\times&\times \end{bmatrix}} \xrightarrow{Q_1} \underset{Q_1A}{\begin{bmatrix} \times&\times&\times\\ 0&\times&\times\\ 0&\times&\times\\ 0&\times&\times\\ 0&\times&\times \end{bmatrix}} \xrightarrow{Q_2} \underset{Q_2Q_1A}{\begin{bmatrix} \times&\times&\times\\ &\times&\times\\ &0&\times\\ &0&\times\\ &0&\times \end{bmatrix}} \xrightarrow{Q_3} \underset{Q_3Q_2Q_1A.}{\begin{bmatrix} \times&\times&\times\\ &\times&\times\\ &&\times\\ &&0\\ &&0 \end{bmatrix}}$$

unchanged

we will not see details. Keep in mind that in Matlab:

$[Q,R] = qr(A,0)$ gives such a factorization.

Note: if A is full column rank (i.e. columns are lin indep) then $\mathrm{Range}(Q) = \mathrm{Range}(A)$.

Unnormalized Simultaneous iteration

Idea: apply power method to many vectors at once.

If $A^k v^{(0)} \to q_1$ we can expect: $\mathrm{span}\{A^k v_1^{(0)}, A^k v_2^{(0)}, \dots, A^k v_n^{(0)}\}$
$\to \mathrm{span}\{q_1, q_2, \dots, q_n\}$

(power method) largest eigenvalue in magnitude

(Simultaneous iteration (~ block power method)) largest n eigenvalues in magnitude

In matrix notation:

$$V^{(0)} = \begin{bmatrix} | & | & & | \\ v_1^{(0)} & v_2^{(0)} & \cdots & v_m^{(0)} \\ | & | & & | \end{bmatrix}$$

$$V^{(k)} = A^k V^{(0)} = \begin{bmatrix} | & | & & | \\ v_1^{(k)} & v_2^{(k)} & \cdots & v_m^{(k)} \\ | & | & & | \end{bmatrix}$$

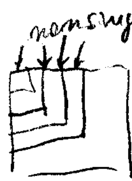
Since all we want is a basis for Range $(V^{(k)})$ (or column space) we use QR facto:

$$\tilde{Q}^{(k)} \tilde{R}^{(k)} = V^{(k)}$$

If we assume:

- $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > |\lambda_{n+1}| \geq \cdots \geq |\lambda_m|$
- all leading principal submatrices of $\tilde{Q}^T V^{(0)}$ are non singular

where $\tilde{Q} = \underbrace{[q_1\ q_2 \cdots q_n]}_{n} \updownarrow m$ = matrix with eigenvectors corresp. to $\lambda_1 \cdots \lambda_n$ as columns.



this condition is similar to condition in power method that initial vector has a component in direction of "largest" eigenvector.

Then it can be shown that:

$$\| q_j^{(k)} - (\pm) q_j \| = O(C^k)$$

where $C = \max_{1 \leq k \leq n} \frac{|\lambda_{k+1}|}{|\lambda_k|}$

(convergence ratio is determined by the eigenvalues that are the closest.)

Simultaneous iteration:

Obviously if we keep multiplying by A, we can have components that quickly become large in magnitude.
$\rightsquigarrow$ as in power method we need to renormalize every time we apply A:

Pick $\tilde{Q}^{(0)} \in \mathbb{R}^{m \times n}$ with orthonormal cols (e.g. from qr facto of a random matrix)

for $k = 1, 2, \ldots$

$$Z = A\,\tilde{Q}^{(k-1)}$$

$$\tilde{Q}^{(k)}\,\tilde{R}^{(k)} = Z$$

It is not hard to show that $\tilde{Q}^{(k)}$ is precisely same matrix for which $\tilde{Q}^{(k)} R = A^k\,\tilde{Q}^{(0)}$ (provided we use same initial guess)

thus convergence study is the same.

QR algorithm (pure version, seldom implemented as is)

$A^{(0)} = A$

for $k = 1, 2, \ldots$

$$Q^{(k)} R^{(k)} = A^{(k-1)}$$

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

factors are recombined in reverse order.

Equivalence of QR algorithm and Simultaneous iteration

Of course this equivalence could only happen if we consider simultaneous iteration on <u>all</u> eigenvectors ($m = n$)

To make equivalence obvious we rewrite algorithms:

Simultaneous iteration	QR algorithm
$\tilde{Q}^{(0)} = I;\ A^{(0)} = A$	$A^{(0)} = A;\ \tilde{Q}^{(0)} = I$
for $k = 1, 2, \dots$	for $k = 1, 2, \dots$
$Z = A\tilde{Q}^{(k-1)}$	$Q^{(k)} R^{(k)} = A^{(k-1)}$
$\tilde{Q}^{(k)} \tilde{R}^{(k)} = Z$	$A^{(k)} = R^{(k)} Q^{(k)}$
$A^{(k)} = \tilde{Q}^{(k)T} A \tilde{Q}^{(k)}$	$\tilde{Q}^{(k)} = Q^{(1)} Q^{(2)} \cdots Q^{(k)}$

We will show by induction that both algorithms generate the same matrices $\tilde{Q}^{(k)}$, $A^{(k)}$ and $\tilde{R}^{(k)} \equiv R^{(k)} R^{(k-1)} \cdots R^{(1)}$, with:

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

ii) $A^{(k)} = (\tilde{Q}^{(k)})^T A \tilde{Q}^{(k)}$

<u>proof</u>: $k=0$ trivial since for both methods $\tilde{Q}^{(0)} = I$, $A^{(0)} = A$, $R^{(0)} = I$

$$I = A^0 = \tilde{Q}^{(0)} \tilde{R}^{(0)} = I\, I.$$

- <u>case $k \geq 1$ for Simult. iter.</u>

$$A^k = A A^{k-1} \underset{\uparrow \text{ induction i)}}{=} \underbrace{A \tilde{Q}^{(k-1)}}\, \tilde{R}^{(k-1)} \underset{\substack{\uparrow \text{ iteration} \\ Z \dots}}{=} \tilde{Q}^{(k)} \underbrace{R^{(k)} \tilde{R}^{(k-1)}}_{= \tilde{R}^{(k)}}.$$

- <u>Case $k \geq 1$ for QR algo</u>

$$A^k = A A^{k-1} \underset{\uparrow \text{ induction i)}}{=} \underbrace{A \tilde{Q}^{(k-1)}}\, \tilde{R}^{(k-1)} \underset{\uparrow \text{ induction ii}}{=} \tilde{Q}^{(k-1)} A^{(k-1)} \tilde{R}^{(k-1)}$$

$$\underset{\text{iteration}}{=} \underbrace{\tilde{Q}^{(k-1)} Q^{(k)}}_{= \tilde{Q}^{(k)}} \underbrace{R^{(k)} \tilde{R}^{(k-1)}}_{= \tilde{R}^{(k)}}$$

Finally to show ii) for QR algo:

$$A^{(k)} \underset{\text{QR algo}}{=} R^{(k)} Q^{(k)} \underset{\text{QR algo}}{=} Q^{(k)T} A^{(k-1)} Q^{(k)}$$

$$\underset{\text{induction}}{=} (Q^{(k)})^T (\tilde{Q}^{(k-1)})^T A \, \tilde{Q}^{(k-1)} Q^{(k)}$$

$$= \tilde{Q}^{(k)T} A \, \tilde{Q}^{(k)}. \qquad \text{qed.}$$

Here is how QR algorithm can be expected to converge.

- QR constructs orthonormal basis of A^k → finds eigenvectors
- diagonal elements of $A^{(k)}$ are Rayleigh quotients:

$$A^{(k)} = (Q^{(k)})^T A \, Q^{(k)}$$

So they should → eigenvalues

- strictly triangular part are "generalized" Rayleigh quotients. Since eigenvectors are ⊥, these should → 0.

We are not done with QR. A more practical version would be as follows: <u>Shifted QR</u>

$A^{(0)} = Q^{(0)} A (Q^{(0)})^T = (\diagdown\!\!\diagdown\!\!\diagdown)$ (Reduction to tridiag form)

for $k = 1, 2, \ldots$

  pick shift $\mu^{(k)}$ (usually $\mu^{(k)} = A^{(k-1)}_{mm}$)

  $Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I$

  $A^{(k)} = R^{(k)} Q^{(k)} + \mu^{(k)} I$

  "deflation": to lock in eigenvalues that have converged