QR Factorisation

Lemma: Given any two vectors $u, v \in \mathbb{R}^n$ with $\|u\|_2 = \|v\|_2$ $\exists w \in \mathbb{R}^n$ s.t. $H(w)u = v$.

Proof: take $w = r(u - v)$, any $r \in \mathbb{R} \setminus \{0\}$, so

$$\begin{aligned} w^T w &= r^2(u^T u - 2v^T u + v^T v) \\ &= 2r^2(u^T u - v^T u) \qquad \text{as } \|u\| = \|v\| \\ &= 2r^2(u - v)^T u = 2rw^T u \end{aligned}$$

so $w^T u = (1/2r)w^T w$. Thus

$$\left(I - \frac{2}{w^T w} w w^T \right) u = u - \frac{2}{w^T w} \frac{w^T w}{2r} \; r(u - v) = v \quad \square$$

In particular
$v = (\|u\|_2, 0, \ldots, 0) = H\left(r\left[u_1 - \|u\|_2, u_2, \ldots, u_n\right]\right) u.$

Can be applied to matrices:

if u is 1^{st} column of A: write $H(w) = H_1, \alpha = \alpha_1$

$$H_1 A = \left[\begin{array}{c|c} \alpha_1 & \times \dots \times \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right] \text{ and if } H(\hat{w}) B = \left[\begin{array}{c|c} \alpha_2 & \times \dots \times \\ \hline 0 & \\ \vdots & C \\ 0 & \end{array}\right]$$

then

$$H_2 = \begin{bmatrix} 1 & 0 \\ 0 & H(\hat{w}) \end{bmatrix} = H\left([0, \hat{w}]\right)$$

satisfies

$$H_2 H_1 A = \left[\begin{array}{c|c|c} \alpha_1 & \times & \times \dots \times \\ \hline 0 & \alpha_2 & \times \dots \times \\ \hline 0 & 0 & \\ \vdots & \vdots & C \\ 0 & 0 & \end{array}\right]$$

continuing inductively for n steps if $m > n$ gives

$$H_n \dots H_2 H_1 A = \begin{bmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_n \\ & & \end{bmatrix} = R \in \mathbb{R}^{m \times n}$$

or for $m-1$ steps if $m \leq n$ gives

$$H_{m-1} \dots H_2 H_1 A = \begin{bmatrix} \alpha_1 & & * \\ & \ddots & \\ 0 & & \alpha_m \end{bmatrix} = R \in \mathbb{R}^{m \times n}$$

Writing
$Q = (H_n \dots H_2 H_1)^{-1} = H_1^T H_2^T \dots H_n^T = H_1 H_2 \dots H_n$
as Householder matrices are symmetric gives

Theorem: Given any $A \in \mathbb{R}^{m \times n}$, $\exists$ an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ and an upper triangular matrix $R \in \mathbb{R}^{m \times n}$ s.t. $A = QR$

Proof: Just take $H_i = I$ if i^{th} column is already zero below diagonal and the above procedure can not break down.

Remark: If $A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \dots & a_n \\ | & | & & | \end{bmatrix}$

and $A = QR$

then $Q = \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \dots & q_n \\ | & | & & | \end{bmatrix}$

and $\{q_1, q_2, \dots, q_n\}$ is an orthonormal basis for span$\{a_1, a_2, \dots, a_n\}$ if this set is linearly independent.

$\Rightarrow QR$ factorization essentially same as Gramm-Schmidt.

Exercise: What happens to QR if $\{a_1, \dots, a_{n-1}\}$ is linearly independent with $a_n \in$ span$\{a_1, \dots, a_{n-1}\}$?

Example: given data $y_1, \ldots, y_m$ at points $x_1, \ldots, x_m$
find parameters in a model e.g. linear model $y = ax + b$
(parameters a, b)
such that

$\sum [y_i - (ax_i + b)]^2$ is min (regression)

same as

$$\min_{a,b} \| \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_m & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \|_2$$

QR Algorithm for eigenvalues (eig in matlab)

Set $A = A_1$

$$\text{for } k = 1, 2, \ldots \quad \begin{cases} \text{factor} & A_k = Q_k R_k & \text{(QR factorisation)} \\ \text{set} & A_{k+1} = R_k Q_k & \text{(matrix multiply)} \end{cases}$$

Lemma: $\{A_k\}$ are all similar matrices and so have same eigenvalues

Proof: $A_{k+1} = R_k Q_k = Q_k^T Q_k R_k Q_k = Q_k^{-1} (A_k) Q_k$ □

Fact: $A_k \to$ upper triangular matrix as $k \to \infty$.

So for large k, $diag(A_k)$ are good approximations to the eigenvalues.

When complex conjugate eigenvalues arise, 2×2 real diagonal blocks remain in A_k which each have a pair of the complex conjugate eigenvalues.

Notes:

1. Speed of convergence depends on the size of gaps between the eigenvalues: more well separated $\Rightarrow$ faster convergence.
2. Convergence is accelerated by the use of shifts (see problem sheet)
3. An orthogonal similarity reduction to Hessenberg form is always employed as a 1^{st} step before apply the QR algorithm with shifts.

Reduction to Hessenberg form

Want a similarity transform $H_p \ldots H_2 H_1 A H_1 H_2 \ldots H_p$ where H_i are Householder matrices but if

$$H_1 A = \left[\begin{array}{c|c} \times & \times \cdots \times \\ \hline \mathbf{0} & \times \cdots \times \\ \vdots & \vdots \quad \vdots \\ \mathbf{0} & \times \cdots \times \end{array}\right]$$

then $H_1 A H_1$ is full i.e. postmultiplication destroys zeros created (otherwise a direct method for eigenvalues!)

So instead let

$$H_1 = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & K_1 \\ 0 & \end{array}\right] \begin{array}{c} 1 \\ \\ n-1 \\ \\ \end{array}, \quad A = \left[\begin{array}{c|c} a_1 & v_1^T \\ \hline u_1 & B_1 \end{array}\right] \text{ then}$$

$$\quad\ 1 \qquad n-1$$

$$H_1 A_1 H_1 =$$

$$\begin{bmatrix} a_1 & v_1^T \\ K_1 u_1 & K_1 B_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & K_1 \end{bmatrix} = \begin{bmatrix} a_1 & v_1^T K_1 \\ K_1 u_1 & K_1 B_1 K_1 \end{bmatrix}$$

and choosing K_1 to be a Householder matrix satisfying $K_1 u_1 = (\alpha_1, 0, \ldots, 0)^T$ we have

$$A^{(2)} = H_1 A H_1 = \begin{bmatrix} a_1 & b_1 & \hat{v}_2^T \\ \alpha_1 & a_2 & v_2^T \\ O & u_2 & B_2 \end{bmatrix} \begin{array}{c} 1 \\ 1 \\ n-2 \end{array}$$

$$1 \quad 1 \quad n-2$$

inductively (similar to before)

$$A^{(n-1)} = H_{n-2} \dots H_2 H_1 A H_1 H_2 \dots H_{n-2}$$

is in Hessenberg form

QR factorization can now be achieved by $n-1$ Givens rotation matrices (see problem sheet) and the Hessenberg form is preserved by the QR algorithm ie. all of the A_k's are upper Hessenberg.