4.3. The QR Reduction

- Reading Trefethen and Bau (1997), Lecture 7
- The **QR** factorization of a matrix $\mathbf{A} \in \Re^{m \times n}$ is

$$\mathbf{A} = \mathbf{QR} \tag{1}$$

  - $\mathbf{Q} \in \Re^{m \times m}$ is an orthogonal matrix
  - $\mathbf{R} \in \Re^{m \times n}$ is upper triangular
  - Assume (for the moment) $m \geq n$
- *Motivation:*
  - An alternate to Gaussian elimination when solving

$$\mathbf{Ax} = \mathbf{b}$$

    * More expensive than Gaussian elimination
  - Determining bases for orthogonal subspaces
  - Determining numerical rank through the SVD
  - Solving least squares problems

$$\min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{b}||_2$$

  - Solving algebraic eigenvalue problems

$$\mathbf{Ax} = \lambda \mathbf{x}$$

Householder QR Factorization

- Use Householder reflections to reduce successive columns of $\mathbf{A}$ to zero below their main diagonals
  - Suppose $m = 6$, $n = 4$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{bmatrix}$$

  - Construct a reflection $\mathbf{H}_1$ such that

$$\mathbf{H}_1\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \\ 0 & a_{52}^{(1)} & a_{53}^{(1)} & a_{54}^{(1)} \\ 0 & a_{62}^{(1)} & a_{63}^{(1)} & a_{64}^{(1)} \end{bmatrix}$$

Householder QR Factorization

- Use a reflection to annihilate $a_{i2}^{(1)}$, $i = 3:6$

$$\mathbf{P}_2 \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ a_{42}^{(1)} \\ a_{52}^{(1)} \\ a_{62}^{(1)} \end{bmatrix} = \begin{bmatrix} r_{22} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

  – Apply the transformation

$$\mathbf{H}_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \end{bmatrix}$$

  to get

$$\mathbf{H}_2\mathbf{H}_1\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & 0 & a_{43}^{(2)} & a_{44}^{(2)} \\ 0 & 0 & a_{53}^{(2)} & a_{54}^{(2)} \\ 0 & 0 & a_{63}^{(2)} & a_{64}^{(2)} \end{bmatrix}$$

Householder QR Factorization

- Repeat the process on the third column

$$\mathbf{H}_3\mathbf{H}_2\mathbf{H}_1\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & a_{44}^{(2)} \\ 0 & 0 & 0 & a_{54}^{(2)} \\ 0 & 0 & 0 & a_{64}^{(2)} \end{bmatrix}$$

- One more should do it

$$\mathbf{H}_4\mathbf{H}_3\mathbf{H}_2\mathbf{H}_1\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ 0 & r_{22} & r_{23} & r_{24} \\ 0 & 0 & r_{33} & r_{34} \\ 0 & 0 & 0 & r_{44} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Householder QR Factorization

- Algorithm for a Householder **QR** factorization

function $\mathbf{A}$ = houseqr($\mathbf{A}$)
% houseqr: Overwrite $\mathbf{A} \in \Re^{m \times n}$ with the product
% $\mathbf{QR}$ where $\mathbf{Q} \in \Re^{m \times m}$ is orthogonal and
% $\mathbf{R} \in \Re^{m \times n}$ is upper triangular. On return,
% the essential part of $\mathbf{Q}$ is stored in the lower triangular
% part of $\mathbf{A}$ and $\mathbf{R}$ is stored in the upper triangular part.

for $j = 1 : n$
    $\mathbf{v}$(j:m) = house($\mathbf{A}$(j:m,j));
    $\mathbf{A}$(j:m,j:n) = row.house($\mathbf{A}$(j:m,j:n), $\mathbf{v}$(j:m));
    if $j < m$
        $\mathbf{A}$(j+1:m,j) = $\mathbf{v}$(j+1:m);
    end
end

Householder QR Factorization

- As described in Example 2.4,

$$\mathbf{H}_j = \mathbf{I} - 2\frac{\mathbf{v}^{(j)}(\mathbf{v}^{(j)})^T}{(\mathbf{v}^{(j)})^T\mathbf{v}^{(j)}} \tag{2a}$$

  – $\mathbf{v}^{(j)}$ is the Householder vector $\mathbf{v}$ used at step j

  – $\mathbf{v}^{(j)}$ has the form

$$\mathbf{v}^{(j)} = [0, 0, ..., 0, 1, v_{j+1}^{(j)}, ..., v_m^{(j)}]^T \tag{2b}$$

- Normalize the Householder vector so that $v_j^{(j)} = 1$

  – This requires modification of the function *house* of Section 4.2. All components of $\mathbf{v}$ are divided by $v(1)$

  – The renormalization saves a storage location and allows $\mathbf{v}^{(j)}$ to be stored in the lower part of $\mathbf{A}$

- Upon completion of the algorithm

$$\mathbf{A} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ v_2^{(1)} & r_{22} & r_{23} & r_{24} \\ v_3^{(1)} & v_3^{(2)} & r_{33} & r_{34} \\ v_4^{(1)} & v_4^{(2)} & v_4^{(3)} & r_{44} \\ v_5^{(1)} & v_5^{(2)} & v_5^{(3)} & v_5^{(4)} \\ v_6^{(1)} & v_6^{(2)} & v_6^{(3)} & v_6^{(4)} \end{bmatrix} \tag{2c}$$

Householder QR Factorization

- $\mathbf{Q}^T = \mathbf{H}_n \mathbf{H}_{n-1} \cdots \mathbf{H}_1$ so

$$\mathbf{Q} = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_n \tag{3}$$

  - $\mathbf{Q}$ rarely needs explicit construction

- The factorization requires about $2n^2(m - n/3)$ FLOPs (multiplications plus additions)
  - With $m = n$ we require about $4n^3/3$ FLOPs
  - Gaussian elimination requires about $2n^3/3$ FLOPs

- With roundoff, we compute

$$\hat{\mathbf{Q}}^T(\mathbf{A} + \delta\mathbf{A}) = \hat{\mathbf{R}}$$

  - $\hat{\mathbf{Q}}$ is *exactly* orthogonal
  - $\hat{\mathbf{R}}$ is upper triangular
  - The perturbation $||\delta\mathbf{A}||_2 \approx u||\mathbf{A}||_2$
  - This is acceptable

Givens QR Factorization

- We can also use Givens rotations to construct the **QR** decomposition
  - The order of the reduction is
  - $\times$ stands for a nonzero entry
  - The numbers indicate the rows that are rotated

$$\mathbf{A} = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{array}{c} \\ \\ 3 \\ 4 \end{array} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ 0 & \times & \times \end{bmatrix} = \begin{array}{c} \\ 2 \\ 3 \\ \\ \end{array} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$$

$$\mathbf{A} = \begin{array}{c} 1 \\ 2 \\ \\ \\ \end{array} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix} = \begin{array}{c} \\ \\ 3 \\ 4 \end{array} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix} = \begin{array}{c} \\ 2 \\ 3 \\ \\ \end{array} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & \times \end{bmatrix}$$

$$\mathbf{A} = \begin{array}{c} \\ \\ 3 \\ 4 \end{array} \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \\ 0 & 0 & 0 \end{bmatrix}$$

Givens QR Factorization

- Algorithm for the Givens **QR** factorization

```
function A = givensqr(A)
% givensqr: Reduce A ∈ ℜ^{m×n} to upper triangular
% form by the Givens QR procedure where Q ∈ ℜ^{m×m}
% is orthogonal and R ∈ ℜ^{m×n} is upper triangular.
% On return, R is stored in the upper triangular part of A.

  for j = 1:n
    for i = m: -1: j + 1
      [c,s] = givens(A(i-1,j), A(i,j));
      A(i-1:i,j:n) = row.rot(A(i-1:i,j:n), c, s);
    end
  end
```

  - The procedure does not save **Q**
  - The factorization requires about $3n^2(m - n/3)$ multiplications and additions
    - Reflections require $2n^2(m - n/3)$ FLOPs
    - Reflections are generally preferred
  - Let $\mathbf{G}_i = \mathbf{G}(i, i-1, \theta)$, then

$$\mathbf{Q} = \underbrace{(\mathbf{G}_m \mathbf{G}_{m-1} \cdots \mathbf{G}_2)}_{1\ st\ col.} \underbrace{(\mathbf{G}_m \cdots \mathbf{G}_3)}_{2\ nd\ col.} \cdots \underbrace{(\mathbf{G}_m)}_{n\ th\ col.}$$

Givens QR Factorization

- *Example 1.* Obtain the **QR** factorization of

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & -3 & 4 \\ 5 & 1 & -1 \\ 7 & 4 & 2 \end{bmatrix}$$

  – Rotating the third and fourth rows using givens(5,7)

$$c = -0.5812, \qquad s = 0.8137$$

  – Applying row.rot to $\mathbf{A}(3:4, 1:3)$ yields

$$\mathbf{A}^{(1)} = \begin{bmatrix} 3.0000 & 2.0000 & 1.0000 \\ 2.0000 & -3.0000 & 4.0000 \\ -8.6023 & -3.8362 & -1.0462 \\ 0 & -1.5112 & -1.9762 \end{bmatrix}$$

  – The Givens matrix corresponding to this rotation is

$$\mathbf{G}(3,4,\theta)^T = \mathbf{G}_4^T$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & -s \\ 0 & 0 & s & c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -0.5812 & -0.8137 \\ 0 & 0 & 0.8137 & -0.5812 \end{bmatrix}$$

  – We may check that $\mathbf{G}_4^T \mathbf{A} = \mathbf{A}^{(1)}$

Givens QR Factorization

- Rotate the second and third rows using givens(2.0000, -8.6023)

$$c = 0.2265, \qquad s = 0.9740$$

  – Applying row.rot to $\mathbf{A}^{(1)}(2:3, 1:3)$ yields

$$\mathbf{A}^{(2)} = \begin{bmatrix} 3.0000 & 2.0000 & 1.0000 \\ 8.8318 & 3.0571 & 1.9249 \\ 0 & -3.7908 & 3.6592 \\ 0 & -1.5112 & -1.9762 \end{bmatrix}$$

  – The rotation matrix is

$$\mathbf{G}_3^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.2265 & -0.9740 & 0 \\ 0 & 0.9740 & 0.2265 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  – We may check that $\mathbf{G}_3^T \mathbf{G}_4^T \mathbf{A} = \mathbf{A}^{(2)}$

Givens QR Factorization

- Rotate the first and second rows using givens(3.0000, 8.8318)

$$c = -0.3216, \qquad s = 0.9469$$

  – Applying row.rot to $\mathbf{A}^{(2)}(1:2, 1:3)$ yields

$$\mathbf{A}^{(3)} = \begin{bmatrix} -9.3274 & -3.5380 & -2.1442 \\ 0 & 0.9104 & 0.3278 \\ 0 & -3.7908 & 3.6592 \\ 0 & -1.5112 & -1.9762 \end{bmatrix}$$

  – The Givens matrix

$$\mathbf{G}_2^T = \begin{bmatrix} -0.3216 & -0.9469 & 0 & 0 \\ 0.9469 & -0.3216 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  – The first column has been reduced

Givens QR Factorization

- Rotate the third and fourth rows using givens(-3.7908, -1.5112)

$$c = 0.9289, \qquad s = -0.3703$$

  – Applying row.rot to $\mathbf{A}^{(3)}(3:4, 2:3)$ yields

$$\mathbf{A}^{(4)} = \begin{bmatrix} -9.3274 & -3.5380 & -2.1442 \\ 0 & 0.9104 & 0.3278 \\ 0 & -4.0809 & 2.6672 \\ 0 & 0 & -3.1908 \end{bmatrix}$$

  – The Givens matrix

$$\mathbf{G}_4^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.9289 & -0.3703 \\ 0 & 0 & 0.3703 & 0.9289 \end{bmatrix}$$

- Rotate the second and third rows using givens(0.9104, -4.0809)

$$c = 0.2177, \qquad s = 0.9760$$

  – Applying row.rot to $\mathbf{A}^{(4)}(2:3, 2:3)$ yields

$$\mathbf{A}^{(5)} = \begin{bmatrix} -9.3274 & -3.5380 & -2.1442 \\ 0 & 4.1812 & -2.5318 \\ 0 & 0 & 0.9007 \\ 0 & 0 & -3.1908 \end{bmatrix}$$

  – The Givens matrix

$$\mathbf{G}_3^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.2177 & -0.9760 & 0 \\ 0 & 0.9760 & 0.2177 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  – The second column is done

The Last Rotation

- Rotate the third and fourth rows using givens(0.9007, -3.1908)

$$c = 0.2717, \qquad s = 0.9624$$

  – Applying row.rot to $\mathbf{A}^{(5)}(3:4,3)$ yields

$$\mathbf{A}^{(6)} = \begin{bmatrix} -9.3274 & -3.5380 & -2.1442 \\ 0 & 4.1812 & -2.5318 \\ 0 & 0 & 3.3154 \\ 0 & 0 & 0 \end{bmatrix}$$

  – The Givens matrix

$$\mathbf{G}_4^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.2717 & -0.9624 \\ 0 & 0 & 0.9624 & 0.2717 \end{bmatrix}$$

The Factored Form

- The factored matrix

$$\mathbf{R} = \begin{bmatrix} -9.3274 & -3.5380 & -2.1442 \\ 0 & 4.1812 & -2.5318 \\ 0 & 0 & 3.3154 \\ 0 & 0 & 0 \end{bmatrix}$$

- The orthogonal matrix
  - We have

$$\mathbf{Q}^T = \underbrace{(\mathbf{G}_4)}_{3\ rd\ col.} \underbrace{(\mathbf{G}_3^T \mathbf{G}_4^T)}_{2\ nd\ col.} \underbrace{(\mathbf{G}_2^T \mathbf{G}_3^T \mathbf{G}_4^T)}_{1\ st\ col.}$$

$$\mathbf{Q} = \begin{bmatrix} -0.3216 & 0.2062 & 0.2511 & 0.8894 \\ -0.2144 & -0.8989 & 0.3813 & 0.0232 \\ -0.5361 & -0.2144 & -0.8121 & 0.0851 \\ -0.7505 & 0.3216 & 0.3635 & -0.4486 \end{bmatrix}$$

  - Check that $\mathbf{QR} = \mathbf{A}$
  - Check that $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$

Properties of the QR Factorization

- **Theorem 1:** Let $\mathbf{A} \in \Re^{m \times n}$ have rank n with $m \geq n$. Let $\mathbf{A} = \mathbf{QR}$, then

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k\}, \qquad k = 1:n \tag{4a}$$

  - $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$ etc.
  - Let $\mathbf{Q}_n = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n]$, $\mathbf{Q}_{m-n} = [\mathbf{q}_{n+1}, \mathbf{q}_{n+2}, \cdots, \mathbf{q}_m]$
  - Then

$$\text{range}(\mathbf{A}) = \text{range}(\mathbf{Q}_n) \tag{4b}$$

$$\text{range}(\mathbf{A})^{\perp} = \text{range}(\mathbf{Q}_{m-n}) \tag{4c}$$

    * The $\perp$ denotes the orthogonal complement
  - In addition

$$\mathbf{A} = \mathbf{Q}_n \mathbf{R}_n, \qquad \mathbf{R}_n = \mathbf{R}(1:n, 1:n) \tag{4d}$$

- *Remark:*

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_n \\ \mathbf{0} \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_n & \mathbf{Q}_{m-n} \end{bmatrix}$$

  - Thus,

$$\mathbf{A} = \mathbf{QR} = \begin{bmatrix} \mathbf{Q}_n & \mathbf{Q}_{m-n} \end{bmatrix} \begin{bmatrix} \mathbf{R}_n \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}_n \mathbf{R}_n$$

- *Proof:* From (1)

$$\mathbf{a}_k = \sum_{j=1}^{k} r_{jk} \mathbf{q}_j \in \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k\} \tag{5}$$

- Thus,

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k\} \subseteq \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k\}$$

Fat and Skinny QRs

- But $\text{rank}(\mathbf{A}) = n$; thus,

$$\dim(\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k\}) = k$$

So

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_k\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k\}$$

- Thus, (4a) is satisfied
- The rest of the Theorem follows directly □
- We call (4c) the *thin* or *reduced* **QR** *factorization* of **A**
- **Theorem 2:** Let $\mathbf{A} \in \Re^{m\times n}$ have rank n, then the thin **QR** factorization (4c) is unique.
  - $\mathbf{R}_n$ is upper triangular with positive diagonal entries
  - Consider the Cholesky factorization of

$$\mathbf{A}^T\mathbf{A} = \mathbf{L}\mathbf{L}^T \tag{6a}$$

    - $\mathbf{L}$ is lower triangular
  - Then

$$\mathbf{R}_n = \mathbf{L}^T \tag{6b}$$

- *Proof:* Consider

$$\mathbf{A}^T\mathbf{A} = \mathbf{R}_n^T\mathbf{Q}_n^T\mathbf{Q}_n\mathbf{R}_n = \mathbf{R}_n^T\mathbf{R}_n$$

  - Comparing this with (6a) yields (6b)
  - The Cholesky factorization is unique
  - From (4d) $\mathbf{Q}_n = \mathbf{A}\mathbf{R}_n^{-1}$; hence, $\mathbf{Q}_n$ is unique □

The Condition Number of a Rectangular Matrix

- What is the condition number of an $m \times n$ matrix?
- **Theorem 3:** Let $\mathbf{A} \in \Re^{n \times n}$, then
  - The eigenvalues λ_i, $i = 1 : n$, of $\mathbf{A}^T\mathbf{A}$ and the singular values σ_i, $i = 1 : n$, satisfy

$$\lambda_i = \sigma_i^2, \qquad i = 1 : n \tag{7a}$$

  - The eigenvectors $\mathbf{x}_i$ of $\mathbf{A}^T\mathbf{A}$ and the singular vectors $\mathbf{v}_i$ of $\mathbf{A}$ satisfy

$$\mathbf{x}_i = \mathbf{v}_i, \qquad i = 1 : n \tag{7b}$$

  - The Euclidean norm

$$||\mathbf{A}||_2 = \sqrt{\rho(\mathbf{A}^T\mathbf{A})} = \sigma_1 \tag{7c}$$

    * The spectral radius $\rho(\mathbf{A}^T\mathbf{A})$ is the largest eigenvalue λ_1 of $\mathbf{A}^T\mathbf{A}$
    * σ_1 is the largest singular value of $\mathbf{A}$
  - The condition number of $\mathbf{A}$ in the Euclidean norm satisfies

$$\kappa_2(\mathbf{A}) = \frac{\sigma_1(\mathbf{A})}{\sigma_n(\mathbf{A})} \tag{7d}$$

The Condition Number of a Rectangular Matrix

- *Proof:*

- We can use this as a definition of the condition number for a rectangular matrix
  - An $O(\epsilon)$ perturbation in $\mathbf{A}$ introduces $O(\epsilon\kappa_2(\mathbf{A}))$ perturbations in $\mathbf{Q}_n$ and $\mathbf{R}_n$
    - cf. Stewart (1993), On the perturbation of LU Cholesky and QR factorizations, *SIAM J. Matrix. Anal.*, **14**, 1141-1145