Lecture 10
Stability of Householder QR and Back Substitution

Songting Luo

Department of Mathematics
Iowa State University

MATH 562 Numerical Analysis II

Outline

1. Stability of Householder QR

2. Stability of Back Substitution

Accuracy of Backward Stable Algorithm

Theorem

If a backward stable algorithm $\tilde{\mathbf{f}}$ *is used to solve a problem* $\mathbf{f}$ *with condition number* κ *using floating-point numbers satisfying the two axioms, then*

$$\|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| / \|\mathbf{f}(\mathbf{x})\| = O(\kappa(\mathbf{x})\epsilon_{machine})$$

Accuracy of Backward Stable Algorithm

Theorem

If a backward stable algorithm $\tilde{\mathbf{f}}$ is used to solve a problem $\mathbf{f}$ with condition number κ using floating-point numbers satisfying the two axioms, then

$$\|\tilde{\mathbf{f}}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\|/\|\mathbf{f}(\mathbf{x})\| = O(\kappa(\mathbf{x})\epsilon_{machine})$$

Proof

Backward stability means $\tilde{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\tilde{\mathbf{x}})$ for $\tilde{\mathbf{x}}$ such that
$\|\tilde{\mathbf{x}} - \mathbf{x}\|/\|\mathbf{x}\| = O(\epsilon_{machine})$
Definition of condition number gives

$$\|\mathbf{f}(\tilde{\mathbf{x}}) - \mathbf{f}(\mathbf{x})\|/\|\mathbf{f}(\mathbf{x})\| \leqslant (\kappa(\mathbf{x}) + o(1))\|\tilde{\mathbf{x}} - \mathbf{x}\|/\|\mathbf{x}\|$$

where $o(1) \to 0$ as $\epsilon_{machine} \to 0$.
Combining the two gives desired result.

Outline

Backward Stability of Householder QR

- For a QR factorization $\mathbf{A} = \mathbf{QR}$ computed by Householder triangularization, the factors $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{R}}$ satisfy

$$\tilde{\mathbf{Q}}\tilde{\mathbf{R}} = \mathbf{A} + \delta\mathbf{A}, \quad \|\delta\mathbf{A}\|/\|\mathbf{A}\| = O(\epsilon_{machine})$$

  i.e., exact QR factorization of a slightly perturbed $\mathbf{A}$ (we will not prove it in class)
- $\tilde{\mathbf{R}}$ is $\mathbf{R}$ computed by algorithm using floating points
- However, $\tilde{\mathbf{Q}}$ is product of exactly unitary reflectors

$$\tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}_1\tilde{\mathbf{Q}}_2\cdots\tilde{\mathbf{Q}}_n$$

  where $\tilde{\mathbf{Q}}_k$ is given by computed $\tilde{\mathbf{v}}_k$, since $\mathbf{Q}$ is not formed explicitly.

Backward Stability of Solving $\mathbf{Ax} = \mathbf{b}$ with QR

Algorithm: solving $\mathbf{Ax} = \mathbf{b}$ by QR factorization

Compute $\mathbf{A} = \mathbf{QR}$ using Householder, represent $\mathbf{Q}$ by reflectors
Compute vector $\mathbf{y} = \mathbf{Q}^*\mathbf{b}$ implicitly using reflectors
Solve upper-triangular system $\mathbf{Rx} = \mathbf{y}$ for $\mathbf{x}$

- All three steps are backward stable
- We will prove for backward substitution later
- Overall, we can show that

$$(\mathbf{A} + \Delta\mathbf{A})\mathbf{x} = \mathbf{b}, \quad \|\Delta\mathbf{A}\|/\|\mathbf{A}\| = O(\epsilon_{machine})$$

as we prove next

Backward Stability of Solving $\mathbf{Ax} = \mathbf{b}$ with Householder QR

Proof: Step 2 gives

$$(\tilde{\mathbf{Q}} + \delta\mathbf{Q})\tilde{\mathbf{y}} = \mathbf{b}, \quad \|\delta\mathbf{Q}\| = O(\epsilon_{machine})$$

Step 3 gives

$$(\tilde{\mathbf{R}} + \delta\mathbf{R})\tilde{\mathbf{x}} = \tilde{\mathbf{y}}, \; \|\delta\mathbf{R}\|/\|\mathbf{R}\| = O(\epsilon_{machine})$$

Therefore,

$$\mathbf{b} = (\tilde{\mathbf{Q}} + \delta\mathbf{Q})(\tilde{\mathbf{R}} + \delta\mathbf{R})\tilde{\mathbf{x}} = [\tilde{\mathbf{Q}}\tilde{\mathbf{R}} + (\delta\mathbf{Q})\tilde{\mathbf{R}} + \tilde{\mathbf{Q}}(\delta\mathbf{R}) + (\delta\mathbf{Q})(\tilde{\mathbf{R}})]\tilde{\mathbf{x}}$$

Step 1 gives

$$\mathbf{b} = [\mathbf{A} + \underbrace{\delta\mathbf{A} + (\delta\mathbf{Q})\tilde{\mathbf{R}} + \tilde{\mathbf{Q}}(\delta\mathbf{R}) + (\delta\mathbf{Q})(\tilde{\mathbf{R}})}_{\Delta\mathbf{A}}]\tilde{\mathbf{x}}$$

where $\tilde{\mathbf{Q}}\tilde{\mathbf{R}} = \mathbf{A} + \delta\mathbf{A}$

Proof of Backward Stability Cont'd

$\tilde{\mathbf{Q}}\tilde{\mathbf{R}} = \mathbf{A} + \delta\mathbf{A}$ where $\|\delta\mathbf{A}\|/\|\mathbf{A}\| = O(\epsilon_{machine})$, and therefore

$$\frac{\|\tilde{\mathbf{R}}\|}{\|\mathbf{A}\|} \leqslant \|\tilde{\mathbf{Q}}^*\| \frac{\|\mathbf{A} + \delta\mathbf{A}\|}{\|\mathbf{A}\|} = O(1)$$

Now show that each term in $\Delta\mathbf{A}$ is small

$$\frac{\|\delta\mathbf{Q}\tilde{\mathbf{R}}\|}{\|\mathbf{A}\|} \leqslant \|\delta\mathbf{Q}\| \frac{\|\tilde{\mathbf{R}}\|}{\|\mathbf{A}\|} = O(\epsilon_{machine})$$

$$\frac{\|\tilde{\mathbf{Q}}\delta\mathbf{R}\|}{\|\mathbf{A}\|} \leqslant \|\tilde{\mathbf{Q}}\| \frac{\|\delta\mathbf{R}\|}{\|\tilde{\mathbf{R}}\|} \frac{\|\tilde{\mathbf{R}}\|}{\|\mathbf{A}\|} = O(\epsilon_{machine})$$

$$\frac{\|\delta\mathbf{Q}\delta\mathbf{R}\|}{\|\mathbf{A}\|} \leqslant \|\delta\tilde{\mathbf{Q}}\| \frac{\|\delta\mathbf{R}\|}{\|\mathbf{A}\|} = O(\epsilon_{machine}^2)$$

Overall

$$\frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} \leqslant \frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\|\delta\mathbf{Q}\tilde{\mathbf{R}}\|}{\|\mathbf{A}\|} + \frac{\|\tilde{\mathbf{Q}}\delta\mathbf{R}\|}{\|\mathbf{A}\|} + \frac{\|\delta\mathbf{Q}\delta\mathbf{R}\|}{\|\mathbf{A}\|} = O(\epsilon_{machine})$$

Since the algorithm is backward stable, it is also accurate.

Outline

1. Stability of Householder QR

2. Stability of Back Substitution

Backward Stability of Back Substitution

- Solve $\mathbf{Rx} = \mathbf{b}$ using back substitution

$$\begin{bmatrix} r_{11} & \cdots & r_{1m} \\ & \ddots & \vdots \\ & & r_{mm} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

- for $j = m$ downto 1

$$x_j = (b_j - \sum_{k=j+1}^{m} x_k r_{jk})/r_{jj}$$

- Back substitute is backward stable

$$(\mathbf{R} + \delta\mathbf{R})\tilde{\mathbf{x}} = \mathbf{b}, \quad \|\delta\mathbf{R}\|/\|\mathbf{R}\| = O(\epsilon_{machine})$$

Furthermore, each component of $\delta\mathbf{R}$ satisfies

$$\frac{|\delta r_{ij}|}{|r_{ij}|} \leqslant m\epsilon_{machine} + O(\epsilon_{machine}^2)$$

- We will show in full detail for $m = 1, 2, 3$ as well as general m

Proof of Backward Stability $(m = 1)$

- For $m = 1$, the algorithm is simply one floating point division. Using floating-point axiom, we get

$$\tilde{x} = b_1 \oslash r_{11} = \frac{b_1}{r_{11}}(1 + \epsilon_1) = \frac{b_1}{r_{11}(1 + \epsilon_1')}$$

  where $|\epsilon_1| \leqslant \epsilon_{machine}$ and $|\epsilon_1'| \leqslant \epsilon_{machine} + O(\epsilon_{machine}^2)$
- Therefore, we solved a perturbed problem exactly:

$$(r_{11} + \delta r_{11})\tilde{x}_1 = b_1, \quad \text{with } \frac{|\delta r_{11}|}{|r_{11}|} \leqslant \epsilon_{machine} + O(\epsilon_{machine}^2)$$

Proof of Backward Stability $(m = 2)$

- For $m = 2$, we first solve $\tilde{x}_2$ as before. Then we compute $\tilde{x}_1$:

$$\tilde{x}_1 = (b_1 \ominus (\tilde{x}_2 \otimes r_{12})) \oslash r_{11} = \frac{b_1 - \tilde{x}_2 r_{12}(1+\epsilon_2))(1+\epsilon_3)}{r_{11}}(1+\epsilon_4)$$
$$= \frac{b_1 - \tilde{x}_2 r_{12}(1+\epsilon_2)}{r_{11}(1+\epsilon_3')(1+\epsilon_4')} = \frac{b_1 - \tilde{x}_2 r_{12}(1+\epsilon_2)}{r_{11}(1+2\epsilon_5)}$$

  where $|\epsilon_2|, |\epsilon_3|, |\epsilon_4| \leqslant \epsilon_{machine}$ and $|\epsilon_3'|, |\epsilon_4'|, |\epsilon_5| \leqslant \epsilon_{machine} + O(\epsilon_{machine}^2)$

- Again, this is an exact solution to $(\mathbf{R} + \delta\mathbf{R})\tilde{\mathbf{x}} = \mathbf{b}$ with

$$\begin{bmatrix} \frac{|\delta r_{11}|}{|r_{11}|} & \frac{|\delta r_{12}|}{|r_{12}|} \\ & \frac{|\delta r_{22}|}{|r_{22}|} \end{bmatrix} = \begin{bmatrix} 2|\epsilon_5| & |\epsilon_2| \\ & |\epsilon_1| \end{bmatrix} \leqslant \begin{bmatrix} 2 & 1 \\ & 1 \end{bmatrix} \epsilon_{machine} + O(\epsilon_{machine}^2)$$

Proof of Backward Stability $(m = 3)$

- For $m = 3$, we first solve $\tilde{x}_3$ and $\tilde{x}_2$ as before. Then we compute $\tilde{x}_1$:

$$\begin{aligned}
\tilde{x}_1 &= (b_1 \ominus (\tilde{x}_2 \otimes r_{12}) \ominus (\tilde{x}_3 \otimes r_{13})) \oslash r_{11} \\
&= \frac{[(b_1 - \tilde{x}_2 r_{12}(1+\epsilon_4))(1+\epsilon_6) - \tilde{x}_3 r_{13}(1+\epsilon_5)](1+\epsilon_7)}{r_{11}(1+\epsilon_8')} \\
&= \frac{b_1 - \tilde{x}_2 r_{12}(1+\epsilon_4) - \tilde{x}_3 r_{13}(1+\epsilon_5)(1+\epsilon_6')}{r_{11}(1+\epsilon_6')(1+\epsilon_7')(1+\epsilon_8')}
\end{aligned}$$

That is, $(\mathbf{R} + \delta\mathbf{R})\tilde{\mathbf{x}} = \mathbf{b}$ with

$$\begin{bmatrix} \frac{|\delta r_{11}|}{|r_{11}|} & \frac{|\delta r_{12}|}{|r_{12}|} & \frac{|\delta r_{13}|}{|r_{13}|} \\ & \frac{|\delta r_{22}|}{|r_{22}|} & \frac{|\delta r_{23}|}{|r_{23}|} \\ & & \frac{|\delta r_{33}|}{|r_{33}|} \end{bmatrix} \leqslant \begin{bmatrix} 3 & 2 & 1 \\ & 2 & 1 \\ & & 1 \end{bmatrix} \epsilon_{machine} + O(\epsilon_{machine}^2)$$

Proof of Backward Stability (general m)

- Similar analysis for general m gives pattern

$$\frac{|\delta \mathbf{R}|}{|\mathbf{R}|} \leqslant \mathbf{W}\epsilon + O(\epsilon_{machine}^2)$$

where $\mathbf{W}$ is (for $m = 5$)

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ & 0 & 1 & 1 & 1 \\ & & 0 & 1 & 1 \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} + \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0 & 1 & 2 & 3 \\ & 3 & 0 & 1 & 2 \\ & & 2 & 0 & 1 \\ & & & 1 & 0 \\ & & & & 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 5 & 1 & 2 & 3 & 4 \\ & 4 & 1 & 2 & 3 \\ & & 3 & 1 & 2 \\ & & & 2 & 1 \\ & & & & 1 \end{bmatrix}$$