Stability and residual

Suppose we have solved a linear system...

```
>> A = randn(4, 4); b = randn(4, 1);
>> x = A \ b;
>> A*x - b
ans =
            0
  -1.3878e-17
            0
   2.2204e-16
```

To check if a computed solution $\widetilde{x}$ is accurate, usually one checks that $A\widetilde{x} - b = r$ ('residual') has small norm.

But does it imply that $\widetilde{x}$ is close to the true solution x?

Resiudal and error

Yes, because $\widetilde{x}$ is the exact solution of

$$A\widetilde{x} = b + r = b + \Delta b$$

and we know that in this case

$$\frac{\|x - \widetilde{x}\|}{\|x\|} \leq \kappa(A)\frac{\|r\|}{\|b\|}.$$

(Note that the computed $r = A\widetilde{x} - b$ might be inaccurate, too, but we are ignoring this issue now.)

Residual test for least squares problems

Now let's turn to least squares problems instead: $\min \|Ax - b\|$, with computed solution $\widetilde{x}$.

Can one expect $r = A\widetilde{x} - b$ to have small norm? No: it's the 'distance' between b and image(A), which can be large (depends on the problem).

What's small then? AF would say 'the derivative: it's a minimum'.

$$A^T A\widetilde{x} - A^T b.$$

Well, kind-of. If $\widetilde{x} = x + \Delta x$, this quantity is $\approx \kappa(A^T A)\|\Delta x\|$. It might be larger

Suppose I have $A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix}$

$$\|Ax - b\| = \|Q^T(Ax - b)\| = \|Q^T A x - Q^T b\| =$$

$$= \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} x - \begin{bmatrix} Q_1^T b \\ Q_2^T b \end{bmatrix} \right\| \quad \boxed{Rx - Q_1^T b = 0}$$

the value of $Ax - b$, computed in the true solution x is $\|Q_2^T b\|$

One could check $\|Ax - b - Q_2^T b\|$ small

or $\|Rx - Q_1^T b\|$ small

$$Rx - Q_1^T b = Q_1^T (Ax - b)$$

So we can check if $Q_1^T(A\tilde{x} - b) \approx 0$

Now this is a "backward stable" check:

$$Q_1^T(A\tilde{x} - b) = r_1$$

Suppose I perturb b to $b + Q_1 r = \tilde{b}$

The solution is given by

$$\min \|Ax - \tilde{b}\| = \|Q^T(Ax - \tilde{b})\| = \left\| \begin{bmatrix} R \\ 0 \end{bmatrix} x - \begin{bmatrix} Q_1^T \tilde{b} \\ Q_2^T \tilde{b} \end{bmatrix} \right\|$$

$$Q_2^T \tilde{b} = Q_2^T (b + Q_1 r) = Q_2^T b + \underbrace{Q_2^T Q_1 r}_{=\,0}$$

because columns of Q_1 and of Q_2 are orthonormal

$$Q_1^T \tilde{b} = Q_1^T (b + Q_1 r) = Q_1^T b + \underbrace{Q_1^T Q_1}_{I} r = Q_1^T b + r$$

If $Q_1^T (A\tilde{x} - b) = r$, then $Q_1^T A\tilde{x}$ solves $Q_1^T A\tilde{x} - b = r$