Lectures on Scientific Computing

Lise-Marie Imbert-Gerard

9/20/18

1 Continuing from last time...(Step 1)

$$\sigma_1 := max_{x \neq 0} \frac{||Ax||}{||x||} > 0$$
$$v_1 \text{such that } ||v_1|| = 1,\ ||Av_1|| = \sigma_1 > 0$$
$$u_1 := \frac{Av_1}{||Av_1||}$$

2 Step 2

$$v_1 = \{x \in \mathbb{R}^m, x^* v_1 = 0\}$$
$$u_1 = \{x \in \mathbb{R}^m, v^* u_1 = 0\}$$
$$A_1 := v_1 \to u_1$$
$$x \to Ax$$

Suppose $A_1 \neq 0$...

$$\sigma_2 := max_{x \in v_1} \frac{||Ax||}{||x||}$$
$$v_2 \text{such that } ||v_2|| = 1, ||Av_2|| = \sigma_2$$
$$u_2 := \frac{Av_2}{||Av_2||}$$

Continue until $A_k = 0$ or $k = m$ or $k = n$. Set all remaining σ_h to and complete $\{u_j\}, \{v_j\}$ into orthogonal bases.

We have build $A = U\Sigma V^*$ so $A^*A = V\Sigma^*\Sigma V^*$. The eigenvalues of A^*A are $\lambda k = \sigma_k^2$ and the eigenvector of A^*A are v_k. We will see that $kl^2 = \frac{\sigma_1(A)}{\sigma_m(A)} \to kl^2(A^*A) = kl^2(A)^2$. This has applications in **Principle Component Analysis** in identifying clusters.

3 Condition Number

$\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Suppose $f\mathbb{K}^n \to \mathbb{K}^n$. Remember that $\Delta f = f(x + \Delta x) - f(x)$.

Definition:

$\mathbb{K}(n) = \lim_{\epsilon \to 0} max_{||\Delta x|| = \epsilon} \frac{||f(x+\Delta x) - f(x)||/||f||}{||\Delta x||/||x||}{}_{\mathbb{R}}$
$R = \frac{||\Delta f||}{||f||} \frac{||x||}{||f||} = \frac{||f'(x)\Delta x||}{||\Delta x||} \frac{||x||}{||f||} + O(||\Delta x||)$
$f'(x)$ is the Jacobian of f at x. Therefore, $\mathbb{K}(x) = ||f'(x)|| \frac{||x||}{||f(x)||}$.

Remark:
If $P, Q \in R^t$ we say that $C > 0$ is a **Sharp Constant** for $P \leq CQ$ if $C' > 0$ such that $C' < C$ and $P \leq C'Q$.

Definition:
We write $P(\epsilon) < CQ(\epsilon)$ as $\epsilon \to 0$ if $P(\epsilon) > 0$, $Q(\epsilon) > 0$, $\lim_{\epsilon \to 0} \frac{P(\epsilon)}{Q(\epsilon)} \leq C$. We say that C is sharp as $\epsilon \to 0$ if it is an identity. The consequence of this is that $K(x)$ is a sharp constant for $\frac{||\Delta f||}{||f||} \leq K(x) \frac{||\Delta x||}{||x||}$ as $||\Delta x|| \to 0$.

3.1 Linear Systems, direct estimates

Let's examine a simple matrix multiplication, i.e. $f(x) = Ax$. Therefore, applying our formula for the condition number $K(x)$

$$K(A, x) = ||A|| \frac{||x||}{||Ax||}$$

The problem is ill-conditioned when K is large, therefore this occurs when $||A|| >> \frac{||Ax||}{||x||}$

3.1.1 Problem

Solving $Au = b$, $u = A^{-1}b$

Applying our formula once again, $K(A^{-1}, b) = ||A^{-1}|| \frac{||b||}{||A^{-1}b}|| = ||A^{-1}|| \frac{||Au||}{||u||}$.

Definition:

$K(A) = ||A|| ||A^{-1}||$ BE CAREFUL HERE! The forward problem and the backward problem do not have the same conditions for stability. This definition can give the mistaken idea that solving the forward (Ax) and backward ($A^{-1}x$) problems have the same stability, but this is not true.

3.1.2 Linear Systems, Perturbation System

What happens when we perturb A and when we perturb b? How does this change the conditioning of the problem? Let's examine the problem $Au = b$ where b is fixed. When we applied perturbation theory previously we found $||\Delta u|| \leq ||A^{-1}||||\Delta A||||u||$. To find the condition number we do some algebra to this expression to yield: $\frac{||\Delta u||}{||u||} \lessapprox ||A^{-1}||||A||\frac{||\Delta A||}{||A||}$. We can check and see that the constant is sharp.

Now let's assume A is fixed and we perturb b. We can do this directly: $\dot{A}u + A\dot{u} = \dot{b} \rightarrow A\Delta u = \Delta b$. We can rewrite this as $\Delta u = A^{-1}\Delta b \rightarrow \frac{||\Delta u||}{||u||} \leq ||A^{-1}||\frac{||\Delta b||}{||b||}\frac{||b||}{||u||}$. We can use the definition $u = A^{-1}b$ or $b = Au$. These provide the inequalities $||u|| \leq ||A^{-1}||||b||$ and $||b|| \leq ||A||||u||$. We want to uses the second definition as it give the value of b bounded by A so we find $\frac{||\Delta u||}{||u||} \leq ||A^{-1}||||A||\frac{||\Delta b||}{||b||}$.

What if neither A or b is fixed? This becomes much more complicated. It combines both definitions and is not that trivial. This is a good case to examine and think about.

3.1.3 Eigenvalue/Eigenvector Problem

$Ar_j = \lambda_j r_j, ||r_j|| = 1$ We want to look at how perturbations of A affect the eigenvalue and eigenvector of this problem.

Looking first at the Eigenvalue...

$$
\begin{aligned}
&Ar_j = \lambda j r_j, A^* = A \\
&r_j^*(\dot{A}r_j + A\dot{r_j} = \dot{\lambda_j}r_j + \lambda_j\dot{r_j}) \\
\therefore\, &r_j^*A\dot{r_j} = (A^*r_j)^*\dot{r_j} = \lambda_j r_j^*\dot{r_j}
\end{aligned}
$$

Since the vector is normalized, $r_j^*r_j = 1$, therefore

$$
\begin{aligned}
&r_j^*\dot{A}r_j = \dot{\lambda_j} \\
&r_j^*\Delta A r_j = \Delta\lambda_j + O(||\Delta A||^2) \\
&|\Delta\lambda_j| \leq ||\Delta A||
\end{aligned}
$$

Therefore, to put it into condition number form,

$$\frac{|\Delta\lambda_j|}{|\lambda_j|} \leq \frac{||\Delta A||}{||A||}\frac{||A||}{\lambda_j||}$$
$$\rightarrow \mathbb{K}_j(A) = \frac{||A||}{|\lambda_j|}$$

If $A^* = A$, $||A||_{l^2} = |\lambda_{max}|$, $\rightarrow$ ill conditioned only if $|\lambda_j| << |\lambda_{max}||$.

Now we want to look at the same problem of eigenvalue in the case where the matrix is not self adjoint ($A \neq A^*$) Since it's not self adjoint we can take that previous inner product with r_j since the values are not necessarily real. But in this case we do know we have left and right eigenvectors! Using this knowledge we know $l_j r_j = 1$ where $l_j A = \lambda_j l_j$. Therefore,

$$l_j(\dot{A} r_j + A\dot{r_j} = \dot{\lambda}_j r_j + \lambda_j \dot{r_j})$$
$$\therefore l_j A\dot{r_j} = \lambda_k l_j \dot{r_j}$$
$$l_j \dot{A} r_j = \dot{\lambda_j} \text{ since } l_j r_j = 1$$
$$\Delta\lambda_j = l_j \Delta A r_j + O(||\Delta A||^2)$$
$$\frac{|\Delta\lambda_j|}{|\lambda_j|} \leq \frac{||l_j^*||||r_j||}{|\lambda_j|}\frac{||\Delta A||}{||A||}||A||$$
$$\rightarrow \mathbb{K}_{\daleth}(A) = \frac{||l_j^*||||r_j||||A||}{\lambda_j}$$
$$||l_j^*||||r_j|| \leq \mathbb{K}_{LinearSystem}(R) = ||R||||R^{-1}||$$

R is the right eigenvector matrix. Ill-conditioning from either $|\lambda_j| << ||A||$ or R is ill-conditioned.

Now let's suppose all eigenvalues are distinct.

Let's assume...
Look this one up.
What we are doing is projecting the basis $\dot{r}$ onto the right eigenvectors