Lecture 8  Condition number of a linear system.

1 Matrix norms

Multiplication by a matrix defines a linear transformation (see Def (6.1) in book). E.g., given a vector \( \mathbf{x} \) and a matrix \( A \), we can compute a transformed vector \( \mathbf{y} = A \mathbf{x} \). Of course, each component of \( \mathbf{y} \) will be transformed differently, but one can ask about some overall magnification (or shrinking) that \( A \) can provide when acting on all possible vectors in a space. This motivates the definition of a following particular case of the definition of a matrix norm (Def 6.19 + (6.22)):

**Def. 1** Let \( \| \cdot \| \) be any of the vector norms (e.g., a \( p \)-norm with \( p = 1, 2, \infty \) — see Lec. 4). Then the corresponding matrix norm induced by this vector norm is:

\[
\|A\| = \max_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} \quad (1)
\]

This is also called a subordinate matrix norm, since it is subordinate to the corresponding vector norm.

**Note 1.** Def. 1 implies that \( \|A\| \geq \frac{\|Ax\|}{\|x\|} \) for any \( \mathbf{x} \), \( \Rightarrow \)

\[
\|Ax\| \leq \|A\| \cdot \|x\| \quad (2)
\]

for any \( \mathbf{x} \).
Note 2. On the other hand, one can show that the maximum in (1) is actually attained, ⇒ there is some \( x_0 \) s.t. \[
\|A\| \cdot \|x_0\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}, \quad ⇒
\]
\[
\|Ax_0\| = \|A\| \cdot \|x_0\| \quad \text{for some } x_0. \quad (3)
\]

\[\text{Note 3 (Key Thm. (6.23))}\]
Let \( A = p \times q \). Then:
\(\text{(a) } \|A\|_1 = \max_j \sum_{i=1}^p |a_{ij}| \) (\text{max. abs. column sum})
\(\text{(b) } \|A\|_\infty = \max_i \sum_{j=1}^q |a_{ij}| \) (\text{max. abs. row sum})
\(\text{(c) } \|A\|_2 = \sqrt{\max \text{ (eigenvalue of } (A^HA))} \)
(\text{In particular, if } A^T = A, \text{ then } \|A\|_2 = \max |A_{ij}|.)

Note 4. Although the matrix norm does give some idea about the magnification that multiplication by \( A \) provides, it does not always give information about the behavior of \( \|A^n\| \) as \( n \to \infty \).

E.g., in Ex. 4 of Sec. 1 (= Ex. 8.17 in book) we had
\[
A = \begin{pmatrix} 0.6 & 0.5 \\ -0.18 & 1.2 \end{pmatrix}
\]
and we saw that \( \|A^n x_0\| \to 0 \) as \( n \to \infty \).
Yet, \[
\|A\|_1 = \max \text{ abs. column sum} = 0.5 + 1.2 = 1.7 > 1
\]
\[
\|A\|_\infty = \max \text{ abs. row sum} = 0.18 + 1.2 = 1.38 > 1.
\]
Thus, based on the fact that \( \|A\| > 1 \), we
cannot conclude that \( \|A^n\| \to \infty \) (or even that \( \|A^n\| \to 0 \)) as \( n \to \infty \).
A formal reason for this will be shown soon.

**Note 5** All subordinate norms are equivalent in the sense that for any two different subordinate norms (say, \( \|A\|_1 \) and \( \|A\|_2 \)), there is \( C_1 \) s.t.
\[
\|A\|_1 \leq C_1 \|A\|_2
\]
for all \( p \times q \) matrices \( A \), where \( C_1 \) depends only on \( p \) & \( q \). (Likewise, there is also \( C_2 = C_2(p,q) \) s.t.
\[
\|A\|_2 \leq C_2 \|A\|_1 .
\]


**Def. 1** is a special case of a more general definition.

**Def. 2** A matrix norm is a function from the vector space \( p \times q \) matrices into the set of nonnegative real numbers s.t.:
1) \( \|A\| > 0 \) and \( \|A\| = 0 \) iff \( A = 0 \) (zero matrix);
2) \( \|aA\| = |a| \|A\| \) for any scalar \( a \);
3) \( \|A + B\| \leq \|A\| + \|B\| \) (triangle inequality);
4) \( \|AB\| \leq \|A\| \|B\| \). \hspace{1cm} (4)

We can now prove that the subordinate norm defined by Def. 1 is indeed a matrix norm. Do 1)-3) at home; will only do 4) here.
Proof of (4)

\[ \| A \times B \| = \max_{x \neq 0} \frac{\| (A(Bx)) \|}{\| x \|} = \max_{x \neq 0} \frac{\| A(Bx) \|}{\| x \|} \leq \max_{x \neq 0} \frac{\| A \| \cdot \| Bx \|}{\| x \|} \]

\[ = \| A \| \cdot \max_{x \neq 0} \frac{\| Bx \|}{\| x \|} \]

\[ \overset{(1)}{=} \| A \| \cdot \| B \| \quad \text{which is (4).} \]

It remains to prove (\( \ast \)).

\[ \| A(Bx) \| \leq \| A \| \cdot \| (Bx) \| \quad \text{by (2),} \]

\[ \Rightarrow \quad (\ast) \quad \text{holds. Thus, (4) is proved.} \]

Note: Later we will give an example of an often-used matrix norm that is not a subordinate norm.

Corollaries of (4)

1. \[ \| A^2 \| \leq \| A \| \cdot \| A \| \]

Similarly, \[ \| A^n \| \leq \| A \| ^n \]

Thus, even though \( \| A \| > 1 \), \( \| A^n \| \to 0 \) as \( n \to \infty \) (see Note 4 after Def. 1).

2. \[ \| A^{-1} \| \geq \frac{1}{\| A \|} \quad (5) \]

for any subordinate norm \( \| \cdot \| \).

(Proof - at home.)

2. Condition number of a matrix and the solution of a perturbed system \( Ax = b \).
In lecture 4 we asked (Q2 & Q9) how sensitive the solution \( x \) of
\[ A \cdot x = b \]  \hspace{1cm} (6)
\[ \] is to a slight change of the entries of \( A \) and \( b \).

**Thm. (a weaker version of Thm. (6.29))**

Let \( A \) be a nonsingular matrix and let \( x \) solve (6). Suppose \( A \) and \( b \) are changed to \((A + \delta A)\) and \((b + \delta b)\), where \( \delta A \) is small in the sense that
\[ \|\delta A \cdot A^{-1}\| < 1 \]  \hspace{1cm} (7)

for some matrix norm. Recall (note 4) that all \( A \) have inverse.

Suppose \((x + \delta x)\) solves
\[ (A + \delta A)(x + \delta x) = b + \delta b \]  \hspace{1cm} (8)

Then the relative change of the solution satisfies:
\[ \|\delta x\| \leq c(A) \cdot \left( \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right) \]  \hspace{1cm} (9)

where
\[ c(A) = \|A\| \cdot \|A^{-1}\| \]  \hspace{1cm} (10)

is called the **condition number** of a matrix.

**Proof:** The general fact under assumption (7) can be derived as a direct superposition of the two subcases considered below (you can do this superposition as a bonus HW problem).
Case I  \[ \Delta A = 0 \]  

Then  
\[ A(x + \Delta x) = \beta + \delta \beta \]  
\[ A\Delta x + A \delta x = \beta + \delta \beta \]  
\[ A\Delta x = \delta \beta \]  
\[ \delta x = A^{-1} \delta \beta \]  
(since \( A \) is nonsingular)  

(2)  
\[ \| \delta x \| \leq \| A^{-1} \| \cdot \| \delta \beta \| . \]  

(11a)  

On the other hand, from (11):  

\[ A\Delta x = \beta \]  
\[ \| \|A\| \| \|\delta x\| \| \geq \| \beta \| \]  
\[ \| \Delta x \| \geq \| \beta \| \]  
\[ \| \beta \| \]  

(11b)  

Divide (11a) by (11b):  

\[ \frac{\| \delta x \|}{\| \beta \|} \leq \frac{\| \|A^{-1}\| \|A\| \| \delta \beta \|}{\| \beta \|} \]  
\[ = c(A) \cdot \frac{\| \delta \beta \|}{\| \beta \|} . \]  

(12)  

Case II  \[ \Delta \beta = 0 \]  

Then  
\[ (A + \delta A)(x + \delta x) = \beta \]  
\[ A\Delta x + \delta A \Delta x + (A + \delta A) \delta x = \beta \]  
\[ (A + \delta A) \delta x = -\delta A \cdot \Delta x \]  
\[ (I + \delta A \cdot A^{-1}) \cdot A \delta x = -\delta A \cdot \Delta x \]  
\[ \approx \]  
\[ A \delta x = -\delta A \cdot \Delta x \]  
\[ \delta x = A^{-1} \cdot \delta A \cdot \Delta x \]  

(2)  
\[ \| \delta x \| \leq \| A^{-1} \| \cdot \| \delta A \| \cdot \| \Delta x \| \]
\[
\frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \left(\frac{\|\delta A\|}{\|A\|}\right).
\] (13)

Doing the calculations in the case \(\delta A \neq 0, \delta b \neq 0\) one can combine (12) & (13) to get (10).

**Discussion** This theorem says that small changes in matrix \(A\) or the r.h.s. \(b\) produce small changes in the solution \(x\) only when \(A\) has a moderate condition number. (In a new problem you'll show that \(\text{cond}(A) > 1\) for any \(A\).) (14)

When \(\text{cond}(A)\) is not "too large", the problem \(A\) is called well-conditioned.

On the other hand, if \(\text{cond}(A) \gg 1\), then small changes in \(A\) or \(b\) may cause very large changes in \(x\), and the solutions of the original problem (6) and perturbed problem (8) may have nothing in common.

Then \(A\) is called ill-conditioned.

In fact, a singular matrix has \(\|A^{-1}\| = \infty\),
\[
\Rightarrow \text{cond}(A) = \|A^{-1}\| \cdot \|A\| = \infty.
\]

Moreover, the condition number shows how close \(A\) is to a singular matrix:
\[
\text{(A+\delta A) is nonsingular as long as } \frac{\|\delta A\|}{\|A\|} < \frac{1}{\text{cond}(A)}.
\] (15)

(see end of p. 272 + Key Thm. (6.28)).
Another way to say this:
"Ill-conditioned" and "nearly singular" are synonyms!

\[ \begin{align*}
\text{Ex.} \quad (10 & \ 7 & 8 & 7) \\
(7 & \ 5 & 6 & 5) \\
(8 & \ 6 & 10 & 9) \\
(7 & \ 5 & 9 & 10)
\end{align*} \begin{bmatrix}
X_1 \\
X_2 \\
X_3 \\
X_4
\end{bmatrix} = 
\begin{bmatrix}
32 \\
23 \\
33 \\
34
\end{bmatrix}
\]

(16)

has the exact solution \[ x = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \]

The solution of (16) with \( \mathbf{b} \) replaced by \( \mathbf{b} + \delta \mathbf{b} = \begin{bmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{bmatrix} \) has \[ x + \delta x = \begin{bmatrix} 9.2 \\ -12.6 \\ 4.5 \\ -1.1 \end{bmatrix}. \] Why??

We can see that \( \det(A) = 1 \), so \( A \) is nonsingular.

But \[ A = \frac{\text{adj}(A)}{\det(A)} \text{ has very large entries,} \]

Lec. 2

\[ \Rightarrow \quad \|A^{-1}\| \text{ is large, and } \Rightarrow \]

\[ C(A) \approx 3000. \]

So a small change of \[ \frac{\|\delta \mathbf{b}\|}{\|\mathbf{b}\|} \approx 0.1 \approx \frac{1}{300} \]

is multiplied by \( C(A) \approx 3000 \) to give changes of the solution of order \[ \frac{\|\delta x\|}{\|x\|} \approx 3000 \cdot \frac{1}{300} \approx 10\%. \]

So, even if \( \det(A) = 1 \) (not close to 0), \( A \) may be close to singular, and its condition \# tells that. Unfortunately, it's impossible to estimate \( C(A) \) w/o a computation (except in some simple cases).