For the three most widely used norms:

\[ \kappa_2 = \text{Cond}_2(A) \]
\[ \kappa_1 = \text{Cond}_1(A) \]
\[ \kappa_\infty = \text{Cond}_\infty(A) \]

**Geometric Interpretation of \( ||| \) , \( \kappa_2 \)**

**Singular Value Decomposition**

It is a proven mathematical identity that any matrix can be written as

\[ A = U \Sigma V^T \]

where, \( U \) and \( V \) are orthogonal matrices and \( \Sigma \) is a diagonal matrix such that \( \sigma_i \geq 0 \).

\[ UU^T = U^T U = I \]
\[ VV^T = V^T V = I \]
\[ \Sigma = \begin{bmatrix}
\sigma_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \sigma_n
\end{bmatrix} \]

Then \( \sigma_i \) are called the singular values of \( A \). Since the matrices \( V \) and \( U \) are orthogonal matrices the magnitude of the matrix \( A \) is contained in the matrix \( \sigma \). The orthogonal matrices are rotating the coordinate system such that \( \sigma \) is the magnitude of matrix along each of the direction of the coordinate system.

\[ A_{m \times n} = U_{m \times m} \begin{bmatrix}
\sigma_1 & \cdots \\
\vdots & \ddots \\
0 & \cdots & \sigma_n
\end{bmatrix} V_{n \times n}^T \]

Assume \( m = n \) (only to simplify the case). In that case

\[ A = U \Sigma V^T \]
\[ \mathbf{AV} = \mathbf{U\Sigma} \]
\[ \mathbf{A}[v_1 \ldots v_i \ldots v_n] = [u_1 \ldots u_i \ldots u_n] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \]

Stable algorithm implies that for a small perturbation in data the numerical result does not suffer large errors. Hence,

\[ |y_{num} - \hat{y}| \text{ is small} \]

where \( \hat{y} \) is the exact result of \( \hat{x} \). Hence a stable algorithm produces nearly the exact result.

\[ |\hat{x} - x| \text{ is small} \]

where \( \hat{x} \) nearly the exact problem.

**Question** How accurate (measure of forward error) is the numerical solution produced by a stable algorithm?

The forward error is given by the expression \( \frac{|y_{num} - y|}{\max(|y_{num}|, |y|)} \). By adding and subtracting \( \hat{y} \) in the numerator and separating the terms we can write

\[ \frac{|y_{num} - y|}{\max(|y_{num}|, |y|)} = \frac{|y_{num} - \hat{y} + \hat{y} - y|}{\max(|y_{num}|, |y|)} \]

An upper bound for the above expression can be found by applying a different denominator for the two terms in the above relation and can be written as

\[ \frac{|y_{num} - y|}{\max(|y_{num}|, |y|)} \leq \frac{|y_{num} - \hat{y}|}{|y_{num}|} + \frac{|\hat{y} - y|}{|y|} \]

\[ \frac{|y_{num} - y|}{\max(|y_{num}|, |y|)} \leq \frac{|y_{num} - \hat{y}|}{|y_{num}|} + \kappa_{\text{problem}} \frac{|\hat{x} - x|}{|x|} \]

Note that an accurate numerical result implies that the value of the forward error is small. This achieved by applying a stable algorithm to a well conditioned problem. There is no guarantee for accuracy if we apply

- Unstable algorithm to an ill-conditioned problem
- Stable algorithm to an ill-conditioned problem
• Unstable algorithm to a well-conditioned problem

It would be interesting to see the outcome of applying an unstable algorithm to an ill-conditioned problem.

Let's look at some common problems and analyze whether they are ill-conditioned or well-conditioned.

Case Study I: Solving Linear Systems

Consider a linear system $Ax = b$, where $\epsilon \ll 1$ and

$$A = \begin{bmatrix} 1 + \epsilon & 1 \\ 1 & 1 \end{bmatrix}$$

$$b = \begin{cases} b_1 \\ b_2 \end{cases}$$

Let's look at the change in the result $x$ for small changes in the system by changing $\epsilon$. The solution of the system $x$ is given by $A^{-1}b$. Consider the case where $\epsilon$ is changed to $2\epsilon$. The change in the system is given by $\|\Delta A\|_\|A\|$ and the change in the solution is given by $\|\Delta x\|_\|x\|$. The change in the system is evaluated as

$$\Delta A = \begin{bmatrix} \epsilon & 0 \\ 0 & 0 \end{bmatrix}$$

Hence the relative change in the norm of the system $A$ is given by

$$\frac{\|\Delta A\|}{\|A\|} \sim \epsilon$$

Similarly, the change in the solution can be evaluated as follows

$$A^{-1} = \frac{1}{\epsilon} \begin{bmatrix} +1 & -1 \\ -1 & 1 + \epsilon \end{bmatrix}$$

$$x = \frac{1}{\epsilon} \begin{cases} b_1 - b_2 \\ -b_1 + (1 + \epsilon)b_2 \end{cases}$$

Hence the relative change in the norm of the solution $x$ is given by,

$$\frac{\|\Delta x\|}{\|x\|} \sim \frac{1}{\epsilon}$$
Figure 1: The transformation of the unit circle by $\mathbf{A}$ into an ellipse

For a small change in the system of the order of $\epsilon$ the solution changes by an amount $\frac{1}{\epsilon}$, which is a very large number.

Consider again the equation

$$\begin{align*}
[\mathbf{A} v_1 \ldots \mathbf{A} v_i \ldots \mathbf{A} v_n] &= [\sigma_1 u_1 \ldots \sigma_i u_i \ldots \sigma_n u_n] \\
\mathbf{A} v_i &= \sigma_i u_i
\end{align*}$$

From linear algebra we know that operating a matrix on a vector produces another vector different length and orientation and that the stretch and the rotation can be separated. As seen in the picture, the matrix $\mathbf{A}$ transforms the unit circle made of two unit vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ into an ellipse of semi axis lengths $\sigma_1$ and $\sigma_2$.

Recall,

$$\| \mathbf{A} \|_2 = \sigma_{\text{max}}$$

$$= \text{Maximum stretch when a vector is operated on by} \mathbf{A}$$

Consider the condition number of the system defined by

$$\kappa_2(\mathbf{A}) = \| \mathbf{A} \|_2 \| \mathbf{A}^{-1} \|_2$$

$$= \sigma_{\text{max}}(\mathbf{A}) \sigma_{\text{max}}(\mathbf{A}^{-1})$$

Let's try to evaluate the maximum SVD of $\mathbf{A}^{-1}$. The matrix $\mathbf{A}$ can be written
in the form
\[
A = U \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} V^T \text{ such that } \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n
\]
Hence, if \( A \) is non-singular then
\[
A^{-1} = V \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_n \end{bmatrix} U^T
\]
So the maximum SVD of \( A^{-1} \) is
\[
\sigma_{\text{max}}(A^{-1}) = \frac{1}{\sigma_{\text{min}}(A)}
\]
Hence the condition number for the system can be evaluated in terms of the SVD values of the matrix \( A \) alone.
\[
\kappa_2(A) = \frac{\sigma_{\text{max}}(A)}{\sigma_{\text{min}}(A)} = \frac{\text{Largest semi axis of the ellipse}}{\text{Smallest semi axis of the ellipse}}
\]
Hence the condition number is a measure of how skewed the ellipse is after transformation.

Let's look at some special cases of the matrix \( A \) and find how the transformed ellipse looks like.

**Case I: A is an Orthogonal Matrix**

Since \( A \) is an orthogonal matrix,
\[
A^T A = I
\]
Writing \( A \) in terms of its singular value decomposition,
\[
I = (U \Sigma V^T)^T U \Sigma V^T \\
= V \Sigma U^T U \Sigma V^T \\
= V \Sigma^2 V^T
\]

5
Figure 2: The matrix $A$ is orthogonal

Pre-multiplying the above relation by $V^T$ and post-multiplying the above relation by $V$, we get

pre-multiplying by $V^T$

$$V^TI = V^T(V\Sigma^2V^T)$$

$$\Rightarrow V^T = \Sigma^2V^T$$

post-multiplying by $V$

$$V^TV = \Sigma^2V^TV$$

$$\Rightarrow I = \Sigma^2$$

$$\Rightarrow \sigma_1 = \sigma_2 = \ldots = \sigma_n = 1$$

Hence all the SVD values of an orthogonal matrix are unity. Hence the condition number of the matrix $\kappa(A)$ is unity. Hence an orthogonal matrix operated on a system of vectors forming a circle is transformed to another unit circle.
Case II: A is a Singular Matrix

In the case when A is a singular matrix, the SVD decomposition is of the form

\[
A = U \begin{bmatrix}
\sigma_1 & & \\
& \ddots & \\
& & \sigma_p \\
& & 0 \\
& & & \ddots \\
& & & & 0
\end{bmatrix} V^T \text{ where } \sigma_1 \geq \sigma_2 \geq \ldots \sigma_p > 0
\]

Hence, the condition number of the system is given by

\[
\kappa_2(A) = \frac{\sigma_{\text{max}}(A)}{\sigma_{\text{min}}(A)} = \frac{\sigma_1}{0} = \infty
\]

Example  Consider the following matrices

\[
U = I, \quad V = I, \quad A = \Sigma = \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}
\]

From the Figure 3 it can be seen that unit circle is transformed into an ellipse which almost reaches a line segment.

Relation between Singular values and Eigen values

The relation is easily seen in an intermediate result of the singular decomposition discussed earlier. If a matrix can be diagonalized then the diagonal matrix contains the eigen values of that matrix. For example consider a matrix A that can be written as

\[
A = X\Lambda X^{-1}
\]

where \( \Lambda \) contains the eigen values of A. Recall the following expression, an intermediary result from singular decomposition.

\[
A^T A = V\Sigma^2 V^T = V\Sigma^2 V^{-1}
\]
Figure 3: The matrix $\mathbf{A}$ is singular

Hence the eigenvalues $\mathbf{A}^T \mathbf{A}$ is contained in the diagonal matrix $\mathbf{\Sigma}^2$. But $\mathbf{\Sigma}$ contains the singular decomposition values of the matrix $\mathbf{A}$. Hence SVD of the matrix $\mathbf{A}$ is the square root of the eigen values of $\mathbf{A} \mathbf{A}^T$.

$$\lambda_i(\mathbf{A} \mathbf{A}^T) = (\sigma_i(\mathbf{A}))^2$$

It should be noted the eigenvalues of $\mathbf{A}$ may be complex or negative real numbers but the singular values of $\mathbf{A}$ are always positive.

**Question**  What is the relation between the singular values, condition number of $\mathbf{A}$ and the impact of data errors in the result?

Consider a two dimensional situation for the ease in physical explanation. Lets look at two situations of a linear system $\mathbf{A} \mathbf{x} = \mathbf{b}$ where $\mathbf{A}$ is orthogonal and singular. The physical representation of the solution in both the cases are shown in Fig. 4 and Fig. 5.

**Case I: $\mathbf{A}$ is Orthogonal**  In both the cases the solid line is the exact line corresponding to the equations of the system. The dotted lines show the errors in the system due to a small change in the vector $\mathbf{b}$. It can be seen from both the figures that a small change in input data $\delta b_1, \delta b_2$ produce a small change and a large change in the output data in case one and two respectively.
Figure 4: The matrix $A$ is orthogonal

Figure 5: The matrix $A$ is singular