

Small change, big effect

Small changes in inputs can cause big changes in the output. For example, the solution set to the linear equations

$$1000x - 1000y = -1000$$

$$1001x - 1000y = -1000$$

is $x = 0, y = 1$. But the solution set to

$$1000x - 1000y = -1000$$

$$1001x - 1000y = -900$$

is $x = 100, y = 101$.

- Decreasing the constant term in the second equation by 10% from -1000 to -900, changes the solution set from $x = 0, y = 1$ to $x = 100, y = 101$.
- It's the butterfly effect.

In matrix form

In matrix form, these two sets of linear equations are

$$\begin{bmatrix} 1000 & -1000 \\ 1001 & -1000 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1000 \\ -1000 \end{bmatrix},$$
$$\begin{bmatrix} 1000 & -1000 \\ 1001 & -1000 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1000 \\ -900 \end{bmatrix}.$$

- These linear equations have the same coefficient matrix and the constant terms differ by at most 10%.
- But the solutions are far apart.
- The determinant of each coefficient matrix 1000.
- A determinant near zero *isn't* required for the butterfly effect.

Consider the linear equations

$$Mx = y, \quad \widehat{M}\widehat{x} = \widehat{y}.$$

- Variables with no hats are the true value, and variables with hats are approximations to the true value.
- We might have, for example,

$$M = \begin{bmatrix} \pi & \sqrt{3} \\ \frac{1}{10} & 23 \end{bmatrix}, \widehat{M} = \begin{bmatrix} \text{fl}(\pi) & \text{fl}(\sqrt{3}) \\ \text{fl}(\frac{1}{10}) & \text{fl}(23) \end{bmatrix}, y = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{42} \end{bmatrix}, \widehat{y} = \begin{bmatrix} \text{fl}(\frac{2}{3}) \\ \text{fl}(\frac{1}{42}) \end{bmatrix},$$

where fl rounds to the nearest float.

- We would like to bound $\|x - \widehat{x}\|$.
- Let's not commit to a specific norm.
- We'll assume that both M and \widehat{M} have inverses.

We know something about $\|y - \hat{y}\|$ and $\|M - \widehat{M}\|$, but we know nothing directly about $M^{-1}y - \widehat{M}^{-1}\hat{y}$; thus

$$\begin{aligned}x - \hat{x} &= M^{-1}y - \widehat{M}^{-1}\hat{y}, \\&= M^{-1}(y - \hat{y} + \hat{y}) - \widehat{M}^{-1}\hat{y}, && \text{(add and subtract)} \\&= M^{-1}(y - \hat{y}) + \left(M^{-1} - \widehat{M}^{-1}\right)\hat{y}. && \text{(regroup)}\end{aligned}$$

So

$$\begin{aligned}\|x - \hat{x}\| &= \|M^{-1}(y - \hat{y}) + \left(M^{-1} - \widehat{M}^{-1}\right)\hat{y}\|, \\&\leq \|M^{-1}(y - \hat{y})\| + \|\left(M^{-1} - \widehat{M}^{-1}\right)\hat{y}\|. \quad \text{(triangle } \leq)\end{aligned}$$

We'll bound each term of the last two terms separately.

We'll assume that for our norm, we have $\|y - \hat{y}\| \leq \varepsilon_m \|y\|$. Thus

$$\begin{aligned}\|M^{-1}(y - \hat{y})\| &\leq \|M^{-1}\| \|y - \hat{y}\|, && \text{(basic norm fact)} \\ &\leq \varepsilon_m \|M^{-1}\| \|y\|, && \text{(assumption)} \\ &= \varepsilon_m \|M^{-1}\| \|Mx\|, && \text{(definition of } x\text{)} \\ &= \varepsilon_m \|M^{-1}\| \|M\| \|x\|. && \text{(basic norm fact)}\end{aligned}$$

On to the second term

Bounding the other term $(M^{-1} - \widehat{M}^{-1})\widehat{y}$ is tricky. A good warm-up is a simple algebra fact:

Warm up For $a, b \in \mathbf{R}_{\neq 0}$, we have

$$a^{-1} - b^{-1} = \frac{1}{a} - \frac{1}{b} = \frac{b - a}{ab} = a^{-1}(b - a)b^{-1}$$

Fact If a and b are invertible matrices, then

$$a^{-1} - b^{-1} = a^{-1}(b - a)b^{-1}.$$

But for matrices a and b , the expression $\frac{b-a}{ab}$ is ambiguous.

$$\begin{aligned}\left(M^{-1} - \widehat{M}^{-1}\right) \widehat{y} &= M^{-1}(\widehat{M} - M)\widehat{M}^{-1}\widehat{y}, \\ &= M^{-1}(\widehat{M} - M)\widehat{x}.\end{aligned}$$

The norm is bounded by

$$\begin{aligned}\left\|\left(M^{-1} - \widehat{M}^{-1}\right) \widehat{y}\right\| &= \|M^{-1}\| \|\widehat{M} - M\| \|\widehat{x}\|, \\ &\leq \varepsilon_m \|M^{-1}\| \|M\| \|\widehat{x}\|.\end{aligned}$$

The product $\|M^{-1}\| \|M\|$ appears in the upper bounds for both of our terms. Let's name it:

$$\text{cond}_M = \|M^{-1}\| \|M\|.$$

We've shown that

$$\|x - \widehat{x}\| \leq \varepsilon_m \text{cond}_M (\|x\| + \|\widehat{x}\|).$$

Our final result

Final refinement:

$$\begin{aligned}\|x - \hat{x}\| &\leq \varepsilon_m \operatorname{cond}_M (\|x\| + \|\hat{x} - x + x\|), \\ &\leq \varepsilon_m \operatorname{cond}_M (\|x\| + \|\hat{x} - x\| + \|x\|).\end{aligned}$$

Combining the $\|x - \hat{x}\|$ terms on the left and the right and assuming $1 - \varepsilon_m \operatorname{cond}_M > 0$ gives

$$\|x - \hat{x}\| \leq \varepsilon_m \frac{2 \operatorname{cond}_M}{1 - \varepsilon_m \operatorname{cond}_M} \|x\|.$$