Chapter 5

Vector and Matrix Norms

5.1 Vector Norms

A vector norm is a measure for the size of a vector.

Definition 5.1. *A* **norm** *on a real or complex vector space V is a mapping $V \to \mathbb{R}$ with properties*

(a) $\|v\| \geq 0 \quad \forall v$

(b) $\|v\| = 0 \quad \Leftrightarrow \quad v = 0$

(c) $\|\alpha v\| = |\alpha| \|v\|$

(d) $\|v + w\| \leq \|v\| + \|w\|$ *(triangle inequality)*

Definition 5.2. *The vector p-norm, $1 \leq p < \infty$, is given by*

$$\|v\|_p = \left(\sum_i |v_i|^p \right)^{1/p} .$$

Special cases:

$$\begin{aligned} \|v\|_1 &= \sum |v_i| \\ \|v\|_2 &= \sqrt{\sum |v_i|^2} \quad \text{(Euclidean norm)} \\ \|v\|_\infty &= \max |v_i|. \end{aligned}$$

The ∞-norm gets its name from

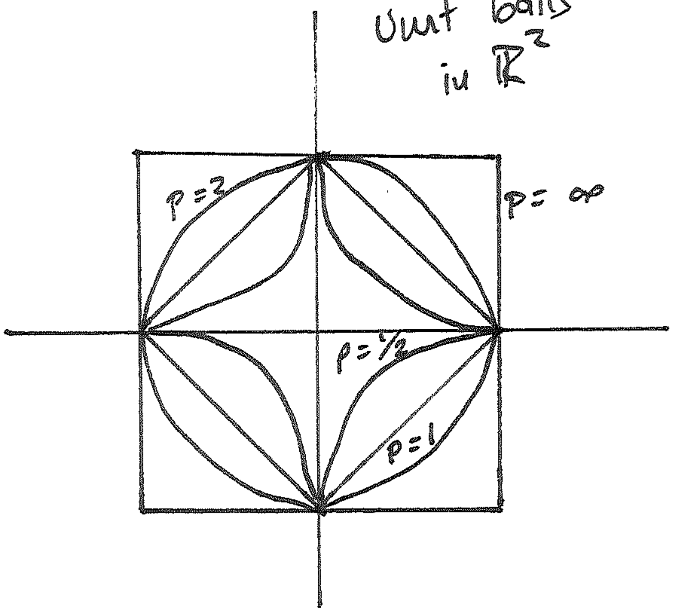
$$\lim_{p \to \infty} \|v\|_p = \|v\|_\infty \quad \forall v.$$

It is easy to verify that conditions (a), (b), (c) are satisfied for all p. The triangle inequality is only satisfied for $p \geq 1$. In fact, it goes the other way for $p < 1$.

Theorem 5.3. *(Hölder Inequality)*

$$|\langle v, w\rangle| \leq \|v\|_p \|w\|_q \qquad \textit{if } \tfrac{1}{p} + \tfrac{1}{q} = 1.$$

The unit balls for these various norms are nested:



Sideline: I defined before

$$\|x\|_0 = \text{number of nonzero entries in } x.$$

The reason for that notation is

$$\lim_{p\to 0} \|x\|_p^p = \|x\|_0.$$

As I just said, for $p < 1$ these are not norms, because the triangle inequality fails. Strangely enough, $\|x\|_0$ satisfies the triangle inequality again, but it is still not a norm because (c) fails.

Slightly more generally, we have **weighted p-norms**

$$\|v\|_{p,w} = \left(\sum_i w_i |v_i|^p \right)^{1/p}.$$

Here the $w_i > 0$ are some fixed weights. This could be useful if some measurements in your data are more reliable than others, or some parts of a solution vector are more important than others.

Theorem 5.4. *If A is positive definite, then*

$$\langle x, y \rangle_A = \langle x, Ay \rangle$$

defines an inner product, and

$$\|x\|_A = \sqrt{\langle x, x \rangle_A}$$

defines a norm.

Example: In the numerical solution of elliptic partial differential equations by finite elements, you can show that the error (difference between numerical solution x_N and true solution x) satisfies

$$\|x - x_N\|_A \leq \quad \text{some estimate,}$$

but you can't get a direct estimate for the standard norm $\|x - x_N\|$.

Here the A is actually a positive definite differential operator, not a matrix, but the idea is the same. ▮

5.1.1 Equivalence of Norms

Sideline: A **relation** R on a set S is a subset of $S \times S$ (ordered pairs):

$$R \subset S \times S = \{(a,b) : a, b, \in S\}.$$

A relation is

- **transitive** if (a,b) and $(b,c) \Rightarrow (a,c)$
- **symmetric** if $(a,b) \Leftrightarrow (b,a)$
- **antisymmetric** if (a,b) and $(b,a) \Rightarrow a = b$.
- **reflexive** if (a,a) for all a.

Example: The ordering $\leq$ on a set of real numbers is transitive, antisymmetric and reflexive. Likewise for the partial set ordering $\subset$. ▮

A relation is called an **equivalence relation** if it is transitive, symmetric and reflexive.

Example: Think of the identity $=$. ▮

Definition 5.5. *Two norms are* **equivalent** *if there are constants* $0 < A \leq B$ *so that*

$$A\|v\| \leq |||v||| \leq B\|v\| \quad \forall v$$

Fact: This is an equivalence relation.
Applications:

- Two equivalent norms have the same notion of convergence. If a sequence converges in one norm, it converges in the other, and vice versa.

- Error estimates are the same, up to a constant factor.

Theorem 5.6. *(Main Theorem in this section) All vector norms in finite dimensions are equivalent.*

Lemma 5.7. *For every norm* $\|\cdot\|$ *on* $\mathbb{C}^n$*, there exists* M *so that*

$$\|v\| \leq M \cdot \|v\|_1.$$

Proof. Let e_i be the ith basis vector. Then $v = \sum v_i e_i$, $\|v\|_1 = \sum |v_i|$, and

$$\|v\| = \left\|\sum_i v_i e_i\right\| \leq \sum |v_i|\|e_i\| \leq M\|v\|_1$$

with $M = \max \|e_i\|$. □

This implies that the norm

$$\|\cdot\| : \mathbb{C}^n \text{ with norm } \|\cdot\|_1 \ \to \mathbb{R}$$

is a continuous function (Lipschitz continuous, actually).

Proof. (of theorem)
We prove that an arbitrary norm is equivalent to the 1-norm.

A set (in a topological space, whatever that is) is **compact** if every open cover has a finite subcover (whatever that means). A major theorem says “A real-valued continuous function on a compact set takes on its maximum and minimum values”.

In $\mathbb{R}^n$ or $\mathbb{C}^n$, compact means the same as “closed and bounded”. The unit sphere under the 1-norm is a closed and bounded set, and by the lemma, the other norm is a continuous function on it. The constants A and B are the maximum and minimum values of this norm on the unit sphere. □

In infinite dimensions, the unit sphere is not compact, and the p-norms are not equivalent.

Example: For the 1, 2, and ∞ norms we have

$$\begin{aligned}\|v\|_2 &\le \|v\|_1 \le \sqrt{n}\|v\|_2\\ \|v\|_\infty &\le \|v\|_2 \le \sqrt{n}\|v\|_\infty\\ \|v\|_\infty &\le \|v\|_1 \le n\|v\|_\infty\end{aligned}$$

I believe some of these inequalities were assigned as a qualifier problem in Summer 2016.

Sample Proof: Let

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Let $\alpha_i = |v_i|/v_i$ if $v_i \neq 0$, $\alpha_1 = 1$ otherwise. Then $|\alpha_i| = 1$, and $\|v_i\| = \alpha_i v_i$. Then

$$\|v\|_1 = \sum_i \alpha_i v_i = \langle v, \bar{\alpha}_i \rangle \le \|v\|_2 \|\bar{\alpha}\|_2 \le \sqrt{n}\|v\|_2.$$

■

Question: If all (finite-dimensional) norms are equivalent, why do we even bother defining different ones?

Answer: Sometimes we can prove estimates for one type of norm much more easily than for another. Also, just because we can.

5.2 The Singular Value Decomposition, Part 1

For any (rectangular) matrix A, the matrix A^*A is square, Hermitian, and positive semidefinite.

Definition 5.8. *The* **singular values** *of A are the square roots of the nonzero eigenvalues of A^*A. It is customary to sort them by size:*

$$\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0.$$

Here r is the rank of A.

Theorem 5.9. *Any matrix A can be factored as*

$$A = U\Sigma V^*,$$

where U, V are unitary and Σ is a diagonal matrix with the σ_i on the diagonal, extended by an appropriate number of 0 if necessary.
This is called the **singular value decomposition** *(SVD).*

If A is of size $m \times n$, then U is $m \times m$, V is $n \times n$, and Σ is $m \times n$.

Facts:

- The σ_i are also the square roots of the nonzero eigenvalues of AA^*. A^*A and AA^* are of different sizes in general, but they have the same nonzero eigenvalues.

- The columns of V are the eigenvectors of A^*A. The columns of U are the eigenvectors of AA^*.

5.3 Matrix Norms

Definition 5.10. *A* **matrix norm** *must satisfy*

(a) $\|A\| \geq 0 \quad \forall A$

(b) $\|A\| = 0 \quad \Leftrightarrow \quad A = 0$

(c) $\|\alpha A\| = |\alpha| \|A\|$

(d) $\|A + B\| \leq \|A\| + \|B\|$

(e) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ *(submultiplicativity)*

Note:

- All the matrix norms we consider are defined for matrices of all sizes. Properties (d) and (e) only apply if the sizes are compatible.

- Some books only require (a)–(d). For me, it does not deserve to be called a matrix norm if it does not satisfy (e) also.

- Notice that (e) implies $\|A^n\| \leq \|A\|^n$. That will be useful later.

- As with vector norms, all matrix norms are equivalent.

▮

Definition 5.11. *A matrix norm and a vector norm are* **compatible** *if*

$$\|Av\| \leq \|A\| \cdot \|v\|$$

This is a desirable property. Note that this definition requires two norms to work together. Typically, a particular matrix norm is compatible with one or more vector norms, but not with all of them.

There are three main sources of matrix norms: (1) vector-based norms; (2) induced matrix norms; (3) norms based on eigenvalues.

We will now look at all of those in turn.

5.3.1 Vector-Based Norms

For a given matrix A, consider the vector vec(A) (the columns of A stacked on top of one another), and apply a standard vector p-norm.

This produces

$$\begin{aligned} p = 1: &\qquad \|A\|_{sum} = \sum_{ij} |a_{ij}| \\ p = 2: &\qquad \|A\|_F = \sqrt{\sum_{ij} |a_{ij}|^2} \\ p = \infty: &\qquad \|A\|_{max} = \max_{ij} |a_{ij}| \end{aligned}$$

The $p = 2$-norm is called the **Frobenius** or **Hilbert-Schmid** norm.

All of them satisfy (a)–(d) automatically. We need to check (e).

Recall two special cases of the Hölder inequality for vector norms:

$$\begin{aligned} |\langle x, y\rangle| &\le \|x\|_2 \cdot \|y\|_2 &\qquad \text{(Cauchy-Schwarz)} \\ |\langle x, y\rangle| &\le \|x\|_1 \cdot \|y\|_\infty & \\ &\le \|x\|_1 \cdot \|y\|_1 &\qquad \text{(obvious)} \end{aligned}$$

Theorem 5.12. *(a) The sum norm satisfies (e)*
(b) The sum norm is compatible with the vector 1-norm.

Proof. (a) Let r_i^* be the ith row of A, c_j the jth column of B. Then

$$\begin{aligned} \|AB\|_{sum} &= \sum_{ij} |(AB)_{ij}| = \sum_{ij} |\langle c_j, r_i\rangle| \\ &\le \sum_{ij} \|r_i\|_1 \cdot \|c_j\|_1 = \|A\|_{sum} \cdot \|B\|_{sum}. \end{aligned}$$

(b) Essentially the same as (a):

$$\|Av\|_1 = \sum_i |(Av)_i| = \sum_i |\langle v, r_i\rangle| \le \sum_i \|r_i\|_1 \cdot \|v\|_1 = \|A\|_{sum} \cdot \|v\|_1.$$

□

Theorem 5.13. *(a) The Frobenius norm satisfies (e)*
(b) The Frobenius norm is compatible with the vector 2-norm.

Proof. Basically the same proof as for the sum norm, except we use Cauchy-Schwarz. □

Lemma 5.14. $\|A\|_F^2 = trace(A^* A)$.

Proof. Write out what trace(A^*A) is, and observe it is equal to $\|A\|_F^2$. □

Theorem 5.15. *If U, V are unitary, then*

$$\|UAV\|_F = \|A\|_F.$$

Proof.

$$\|UA\|_F^2 = \text{trace}((UA)^*(UA)) = \text{trace}(A^*U^*UA) = \text{trace}(A^*A) = \|A\|_F^2.$$

Similarly for V. □

There will be more properties of the Frobenius norm in section 5.3.3.
Fact: The max-norm does not satisfy (e).
Exercise: Find a counterexample.

5.3.2 Induced Matrix Norms

Definition 5.16. *Given any vector norm, the* **induced matrix norm** *is given by*

$$\|A\| = \sup_{v \neq 0} \frac{\|Av\|}{\|v\|} = \sup_{\|v\|=1} \|Av\|.$$

It is easy to check that (a)–(e) are satisfied, and that these norms are automatically compatible with the vector norm that produced them.

Theorem 5.17.

$$\|A\|_1 = \max_j \sum_i |a_{ij}| \qquad \textit{(largest column sum)}$$

$$\|A\|_\infty = \max_i \sum_j |a_{ij}| \qquad \textit{(largest row sum)}$$

$$\|A\|_2 = \textit{largest singular value}$$

Proof.

$$\begin{aligned}
\|Av\|_1 &= \sum_i |(Av)_i| \leq \sum_i \sum_j |a_{ij}| \cdot |v_j| \\
&= \sum_j \left(\sum_i |a_{ij}| \right) |v_j| \leq \sum_j \left(\max_k \sum_i |a_{ik}| \right) \cdot |v_j| \\
&= \left(\max_k \sum_i |a_{ik}| \right) \cdot \|v\|_1.
\end{aligned}$$

This proves that

$$\|A\|_1 \leq \max_j \sum_i |a_{ij}|.$$

To complete the proof, we need to find one particular v for which we get equality. Assume that the largest column sum is in column j_0, then $v = e_{j_0}$ (standard basis vector) will work.

The proof for $p = \infty$ is similar (exercise).

The proof for $p = 2$ will be done later, in corollary 5.21. □

Example: Let

$$A = \begin{pmatrix} 3 & -1 & 4 \\ 1 & 5 & -9 \\ 2 & 6 & 5i \end{pmatrix}.$$

The row sums are 8, 15, 13. The column sums are 6, 12, 18.

$$\|A\|_1 = 18, \qquad \|A\|_\infty = 15, \qquad \|A\|_2 \approx 13.5824.$$

▮

Induced Norms of Special Matrices

For a few types of matrices, some of the induced matrix norms are easy to calculate.

Theorem 5.18. *If*

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

is a diagonal matrix, then $\|D\|_p = \max_i |d_i|$ *for all* $p \geq 1$.

Proof.

$$Dv = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} d_1 v_1 \\ \vdots \\ d_n v_n \end{pmatrix}.$$

Then

$$\|Dv\|_p^p = \sum_i |d_i|^p |v_i|^p \leq \left(\max_i |d_i|^p \right) \sum_i |v_i|^p = \left(\max_i |d_i| \right)^p \|v\|_p^p,$$

so

$$\|Dv\|_p \leq \max_i |d_i|.$$

To show equality, you need to find one particular vector for which you have equality. For example, the standard basis vector e_{i_0} for the index i_0 which corresponds to the maximum $|d_i|$. □

Theorem 5.19. *If U is unitary, then* $\|U\|_2 = 1$.

Proof. $\|Uv\|_2^2 = \langle Uv, Uv \rangle = \langle v, U^* Uv \rangle = \langle v, v \rangle = \|v\|_2^2$. □

Theorem 5.20. *If U, V are unitary, then $\|UAV\|_2 = \|A\|_2$.*

Proof.

$$\begin{aligned}\|UA\|_2 &\le \|U\|_2\|A\|_2 = \|A\|_2,\\ \|A\|_2 &= \|U^*UA\|_2 \le \|U^*\|_2\|UA\|_2 = \|UA\|_2\end{aligned}$$

Likewise for V. □

Corollary 5.21. *$\|A\|_2$ = the largest singular value.*

Theorem 5.22. *If U, V are unitary, then $\|UAV\|_F = \|A\|_F$.*

Proof. Consider the singular value decomposition

$$A = U\Sigma V^*.$$

By theorem 5.20, $\|A\|_2 = \|\Sigma\|_2$. By theorem 5.18, $\|\Sigma\|_2 = \sigma_1$. □

5.3.3 Matrix Norms Based on Eigenvalues

There cannot be any norms based on the eigenvalues of A itelf, because there are non-zero matrices with only zero eigenvalues, for example

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Instead, you need to base the norms on the singular values. We will get to that in a bit. First, we prove a few theorems about eigenvalues and norms in general.

The Spectral Radius and Norms

Theorem 5.23. *$\rho(A) \le \|A\|$ for any matrix norm.*

Proof. If $\|A\|$ is compatible with any vector norm, this is easy: Take v to be the eigenvector to the largest eigenvalue, then

$$|\lambda|\|v\| = \|Av\| \le \|A\|\|v\|.$$

For a general norm (which satisfies (e)) we use a littletrick: Let V be the matrix all of whose columns are equal to v. Then

$$|\lambda|\|V\| = \|AV\| \le \|A\|\|V\|.$$

□

Theorem 5.24. *For any (square) matrix A and any $\epsilon > 0$ there exists a matrix norm so that*

$$\rho(A) \leq \|A\| \leq \rho(A) + \epsilon.$$

Note: This does not say that there is a single matrix norm that works for all matrices A. It says that for each fixed A and fixed ϵ, there is such a norm. ▮

Corollary 5.25. *If $\rho(A) < 1$, then $A^n \to 0$.*

Proof. Let $\rho(A) = 1 - \epsilon$, and find a matrix norm so that $\|A\| < 1 - (\epsilon/2) < 1$. Then $\|A^n\| \leq \|A\|^n \to 0$. □

Theorem 5.26. *For any matrix norm,*

$$\rho(A) = \lim_{n \to \infty} \|A^n\|^{1/n}.$$

Proof. $\rho(A^n) = \rho(A)^n \leq \|A^n\|$, so

$$\rho(A) \leq \|A^n\|^{1/n}$$

for all n, so therefore also in the limit.

For the opposite direction, choose any $\epsilon > 0$. Let

$$A_\epsilon = \frac{1}{\rho(A) + \epsilon} A,$$

then

$$\rho(A_\epsilon) = \frac{\rho(A)}{\rho(A) + \epsilon} < 1.$$

There is some matrix norm for which $\|A_\epsilon\| < 1$, so $\|A_\epsilon^n\| \to 0$. Since all matrix norms are equivalent, this also applies to whatever matrix norm we are dealing with. For large enough n, $\|A_\epsilon^n\| < 1$, which implies

$$\|A^n\|^{1/n} \leq \rho(A) + \epsilon.$$

□

Remark: If a matrix A has spectral radius ρ, can $\|Av\|$ every be larger than $\rho \cdot \|v\|$? The answer is yes, for example

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \rho(A) = 1.$$

What the theorem says is that in the long run, if you keep applying A over and over again, on average the vector cannot grow by any factor larger than ρ. ▮

Sideline: A concept related to the spectral radius is the **numerical radius** of a matrix. The **numerical range** of A is

$$W(A) = \{\langle Av, v\rangle : \langle v, v\rangle = 1\}.$$

This is a subset of the complex plane which includes the spectrum. The numerical radius is

$$r(A) = \max_{z \in W(A)} |z|.$$

Obviously, $r(A) \geq \rho(A)$.

The numerical radius does better than the spectral radius: it satisfies conditions (a)–(d) of a matrix norm, just not (e). A counterexample is

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

with

$$r(A) = r(B) = 1/2, \qquad r(AB) = 1.$$

The Schatten Norms

Definition 5.27. *The* **Schatten** *p-norm of A is the vector p-norm applied to the vector of singular values*

$$\sigma = (\sigma_1, \cdots, \sigma_r)^T.$$

As usual, the interesting cases are $p = 1$, $p = 2$, $p = \infty$.

$$\begin{aligned} p = 1: &\qquad \|A\|_* = \sum_i \sigma_i \qquad \text{(the } \textbf{nuclear norm}\text{)} \\ p = 2: &\qquad \|A\|_F = \sqrt{\sum_i \sigma_i^2} \\ p = \infty: &\qquad \|A\|_2 = \max_i \sigma_i \end{aligned}$$

The nuclear norm is new. The Schatten 2-norm turns out to be the Frobenius norm. The Schatten ∞-norm is the matrix 2-norm.

Proof. We have $A = U\Sigma V^*$ (SVD), and for both Frobenius norm and 2-norm we proved earlier that $\|A\| = \|\Sigma\|$. The rest is then obvious. □

Sideline: It is interesting to consider what the Schatten 0-norm (which is not really a norm) would be. This is the number of nonzero singular values, which is the rank of A.

5.4 Applications of Matrix Norms

5.4.1 Fun with Matrix Power Series

Recall from calculus a few facts about power series. We will only consider power series about 0.

- A **sequence** $\{x_0, x_1, \ldots\}$ converges to a limit L if $|x_n - L| \to 0$ as $n \to \infty$. You can phrase that in terms of epsilons and deltas yourself.
- A **power series** is an infinite series $\sum_{k=0}^{\infty} a_k x^k$. It **converges** if the sequence of partial sums $\sum_{k=0}^{n} a_k x^k$ converges. It **converges absolutely** if $\sum_{k=0}^{n} |a_k||x|^k$ converges.
- Every power series has a **radius of convergence** R. If $|x| < R$, the series converges absolutely, if $|x| > R$, it diverges. If $|x| = R$, anything can happen. Here x can be real or complex.

We can also consider power series of matrices, of the form $\sum_{k=0}^{\infty} a_k A^k$. Because of $\|A^k\| \leq \|A\|^k$, we immediately get

Lemma 5.28. *If $\|A\| < R$ for any matrix norm, or if $\rho(A) < R$, then the power series converges.*

Note: If $\rho(A) > R$, the series will diverge. If $\|A\| > R$ for some norm, that is inconclusive, since there may be some other norm which is less than R. ▮

Example: The power series for $1/(1-x)$ is

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k.$$

The radius of convergence is 1.

If $\rho(A) < 1$, then $I - A$ is invertible, and

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

If $\|A\| < 1$ in some norm, then in the same norm

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

▮

Sideline: If $x(s)$ is a function on some interval $[a, b]$,

$$x(s) = f(s) + \int_a^b k(s,t)x(t)\,dt$$

is called a **Fredholm integral equation**. f is a known function, and k is called the **kernel**.

The discrete counterpart is the linear equation $x = f + Kx$. By the previous example, if $\rho(K) < 1$, this has a unique solution x for every f, and x depends continuously on f.

Similar statements are true for the original integral equation.

5.4.2 The Condition Number

This is a topic from numerical analysis.

Consider a mathematical problem with input x, output y. We calculate y from x in some fashion: $y = F(x)$.

Example: Find the zeros of a polynomial

$$p(t) = a_n t^n + \cdots + a_1 t + a_0.$$

Here the input x is the coefficient vector $(a_n, \ldots, a_0)^T$, and the output y is the vector of zeros $(t_1, \ldots, t_n)$. ▮

If we change the input a little, from x to $x + \Delta x$, we get a different output $y + \Delta y$. Δx could be measurement error, or roundoff error from putting numbers on a computer. The question is: How sensitive is the output to small changes in input?

Definition 5.29. *The* **absolute error** *is* $\|\Delta x\|$ *or* $\|\Delta y\|$. *The* **relative error** *is* $\|\Delta x\|/\|x\|$ *or* $\|\Delta y\|/\|y\|$.

Usually, the relative error is more meaningful. A relative error of 10^{-6} means that we can trust 6 decimals in the number.

Definition 5.30. *The* **magnification factor** *for the (relative) error is*

$$\frac{\|\Delta x\|/\|x\|}{\|\Delta y\|/\|y\|}.$$

This magnification factor depends on Δx.

The **condition number** *of the problem is the worst possible case of error magnification:*

$$cond = \max_{small\ \Delta x} \frac{\|\Delta x\|/\|x\|}{\|\Delta y\|/\|y\|}.$$

This requires some comments:

- The condition number depends on the norm used. Different norms give different condition numbers, but usually of the same order of magnitude.

- The condition number is an interesting concept, but most of the time you cannot actually calculate it. One exception is in linear algebra.

- I am being deliberately vague about the meaning of "small Δx". In linear algebra, it does not really matter, since constant multiples cancel out. You can just put "any Δx" there.

- A problem with large condition number is called **ill-posed**. This concept was introduced by Hadamard, who argued than anything coming up in real life had to be well-posed (small condition number). He was wrong. There are many problems that are ill-posed, but nevertheless can provide meaningful results. Examples include CAT scans, and weather prediction. The trick is not to require too much accuracy.

- There is also a difference between the condition number of the problem (assuming you can find the mathematically exact solution), and the condition number of a particular algorithm on the computer. If the problem itself is ill-posed, no algorithm can fix that, but it is possible to have an algorithm that has a worse condition number than the underlying problem. Don't use that one.

Example: This example explains how you would use the condition number.

Suppose you are calculating something on a standard computer. The computer works with fixed accuracy, usually equivalent to about 15 decimals of accuracy.

Suppose you know that the condition number of your problem is 10^{10}.

Whenever you put your numbers on the computer, you have to assume a $\|\Delta x\|$ of at least 10^{-15}, because your input gets rounded to 15 decimals. In the worst case, that error will be magnified by 10^{10}, so your final relative error could be 10^{-5}. That means you can only trust 5 decimals in your answer.

Caution: This is assuming the worst case, both in the original rounding and in the error propagation. Most of the time, your answer will be correct to more decimals. The point is that you cannot **trust** any more than 5 decimals. ▮

Let us consider one special case: solving a system of linear equations $Ax = b$. We will only consider the condition number for changes in the right-hand side b.

So: A is fixed, the input is b, and the output is x. We consider

$$\begin{aligned} Ax &= b \\ A(x + \Delta x) &= (b + \Delta b) \end{aligned}$$

which implies

$$A\Delta x = \Delta b \qquad \Rightarrow \qquad \Delta x = A^{-1}\Delta b.$$

The condition number is

$$\kappa = \max_{\Delta b} \frac{\|\Delta x\|/\|x\|}{\|\Delta b\|/\|b\|}.$$

Now

$$\|b\| \leq \|A\| \cdot \|x\| \qquad \Rightarrow \qquad \|x\| \geq \frac{\|b\|}{\|A\|}, \qquad \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

$$\|\Delta x\| \leq \|A^{-1}\| \|\Delta b\|.$$

Together we find that

$$\kappa \leq \|A\| \cdot \|A^{-1}\|.$$

It is possible to actually find specific a Δb which achieves this bound (exercise for the reader), so

$$\kappa = \|A\| \cdot \|A^{-1}\|.$$

Comments:

- This is the condition number for changes in the right-hand side b. You can also consider the condition number for changes in A. That turns out the be the same number, but **that is a coincidence**. If you consider the least squares solution of an overdetermined $Ax = b$, the condition numbers for changes in b and changes in A are different.
- The condition number for the forward problem "compute $y = Ax$" also happens to be the same.
- As I said above, for different norms you get different condition numbers. For the 2-norm, $\kappa = \sigma_1/\sigma_n$ (ratio of largest and smallest singular value).
- In theoretical linear algebra, a matrix is either singular, or it is not. In practical linear algebra, there is no such thing as an exactly singular matrix, unless you have a matrix of small integers. What happens is that $\|A^{-1}\|$, and therefore κ, gets so large that you have no digits of accuracy left, so you computations are basically meaningless.