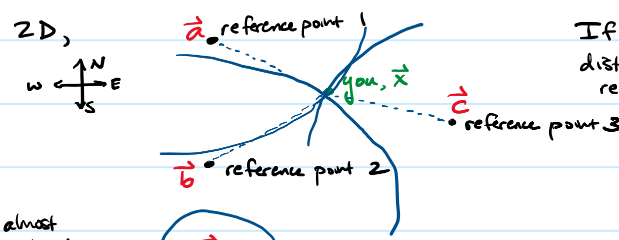
Ch 10.1: Fixed Points for Multivariate Functions and Contraction Mapping Theorem

Tuesday, September 16, 2025 2:14 PM

Motivation (not in book)

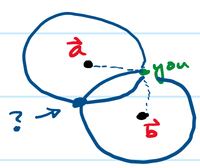
How does the Global Positioning System (GPS) work?

Triangulation: in 2D,



If you know the distance to 3 known reference points, you can figure out your location

2 points is almost enough but there's a slight ambiguity



In 3D, we need 4 reference points: $\vec{a}, \vec{b}, \vec{c}, \vec{d} \in \mathbb{R}^3$ ← known locations

and your distances to each: $\|\vec{x}-\vec{a}\|_2, \ldots, \|\vec{x}-\vec{d}\|_2$

For GPS, reference points are satellites, and their location at any given time is known precisely.

How do we measure our distance to a satellite, $\|\vec{x}-\vec{a}\|_2$?

- Satellites send out a message that includes information about when they sent it (very precise! They have atomic clocks)
- You receive it, and can compute how long it took to travel from the satellite to you.
- The distance is proportional to the travel time ✓
  - i.e. proportionality constant is speed of light.

But there's a catch!
Your GPS receiver (or cell phone) doesn't have an atomic clock.
You don't know the time (not to high enough accuracy).

Solution: solve for the time! 4 GPS satellites is (almost) redundant. Use extra info.

i.e. solve the following nonlinear system of equations:

4 equations
$$\left\{\begin{aligned}
\|\underset{\text{unknown}}{\vec{x}} - \underset{\text{known}}{\vec{a}}\|_2^2 &= s^2 \cdot (\underset{\text{unknown}}{t} - \underset{\text{known}}{t_a})^2 \\
\|\vec{x}-\vec{b}\|_2^2 &= s^2 \cdot (t-t_b)^2 \\
\|\vec{x}-\vec{c}\|_2^2 &= s^2 \cdot (t-t_c)^2 \\
\|\vec{x}-\vec{d}\|_2^2 &= s^2 \, (t-t_d)^2
\end{aligned}\right.$$

"distance = rate × time" or $d^2 = \text{rate}^2 \times t^2$

"s" = speed of light (since "c" is already in use)

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \; t \;\Big\}$ 4 unknowns $\qquad \|\vec{x}-\vec{a}\|_2^2 = \sum_{i=1}^{3} (x_i - a_i)^2$ distance formula

... your Phone GPS solves these nonlinear equations to use google maps, etc. !!

Ch 10.1, p. 2

Tuesday, September 16, 2025 2:37 PM

Problems we'll solve in ch. 10

- Solve multivariate equations*, $F(\vec{x}) = \vec{0}$, $F(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{bmatrix}$

  "root finding"

  $\vec{x} \in \mathbb{R}^n$

  $F: \mathbb{R}^n \to \mathbb{R}^n$ for now n eq'n, n unknowns

  * if F is linear, use specialized linear algebra techniques instead

  Assume F is differentiable

  (that's why we did $d^2 = r^2 t^2$ not $\underbrace{d = rt}_{\text{non-smooth}}$ on GPS example)

- solve multivariate fixed point problems: find $\vec{p} \in \mathbb{R}^n$ st. $G(\vec{p}) = \vec{p}$

  obviously related to $F(\vec{x}) = 0$, eg. $F(\vec{x}) = G(\vec{x}) - \vec{x}$

- optimization, $\min f(\vec{x})$, $f: \mathbb{R}^n \to \mathbb{R}$

  related to root finding since we want $\overbrace{\nabla f(\vec{x}) = \vec{0}}^{\text{stationary point}}$

  More subtle though, since 1) we distinguish saddle pts vs. local min vs. global min

  i.e. not all stationary pts. are equally good

  These issues are beyond the scope of this class

  2) constraints in $\mathbb{R}^n$ are not as simple as "checking the endpoints"

Fixed point equations

Basic existence question: do we have anything like Thm 2.3 (i)? like IVT?

Naive extension of IVT to 2D (or n-dim) fails.

Answer: famous Brouwer Fixed Pt. thm (our book's Thm 10.6 part 1 is a special case)

Thm (Brouwer, 1909)   Nontrivial proof... can use "hairy ball thm"

Let $D \subseteq \mathbb{R}^n$ be a nonempty *compact convex set and let

$G: D \to \mathbb{R}^n$ be ① continuous

② $\forall \vec{x} \in D$, $G(\vec{x}) \in D$ "G maps D into D"

then G has at least one fixed point in D, i.e. $\exists \vec{p} \in D$ such that $\vec{p} = G(\vec{p})$

*For example, a closed ball $D = \{ \vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| \le r \}$

or a hyperrectangle $D = \left\{ \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : a_i \le x_i \le b_i,\ i = 1, \dots, n \right\}$ ← Book specializes to this case

Tuesday, September 16, 2025 2:58 PM

Finding fixed points

We'll generalize what we did in 1D

p is unrelated to n. we choose $1 \le p \le \infty$

Def $G: \mathbb{R}^n \to \mathbb{R}^n$ is *Lipschitz continuous* with respect to the p-norm if $\exists L$ st. $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$, $\|G(\vec{x}) - G(\vec{y})\|_p \le L \cdot \|\vec{x} - \vec{p}\|_p$

Also unrelated to our fixed point $\vec{p}$

or if $G: D \to \mathbb{R}^n$, we change $\forall \vec{x}, \vec{y} \in \mathbb{R}^n$ to "$\forall \vec{x}, \vec{y} \in D$"

Def $G: \mathbb{R}^n \to \mathbb{R}^n$ or $G: D \to \mathbb{R}^n$ is *contractive* with respect to the p-norm (or any norm) if it has a Lipschitz constant less than 1 (in that norm)

* really, any norm works. $\|\cdot\|$ is a norm on $\mathbb{R}^n$ as long as it satisfies: (maybe skip in class?)

1) $\|\vec{x}\| \ge 0$, $\|\vec{x}\| > 0$ except $\|\vec{0}\| = 0$
2) $\|\alpha \cdot \vec{x}\| = |\alpha| \cdot \|\vec{x}\|$
3) $\|\vec{x} + \vec{y}\| \le \|\vec{x}\| + \|\vec{y}\|$ "triangle ineq."

Common norms: $\|\vec{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ including $\underbrace{\|\vec{x}\|_1 = \sum |x_i|}_{\text{time it takes in Manhattan (on a grid)}}$, $\underbrace{\|\vec{x}\|_2 = \sqrt{\sum x_i^2}}_{\text{time as the crow flies}}$ "Euclidean norm"

$\underbrace{\|\vec{x}\|_\infty = \max_{i=1,\dots,n}}_{\text{time it takes to leave for a party if you have to wait for everyone}} |x_i|$

Thm (similar to 2nd part of Thm 10.6) "Contraction Mapping" aka "Banach-Picard"

Let $G: D \to \mathbb{R}^n$ where $D \subseteq \mathbb{R}^n$ is a *closed set, and suppose G is ① contractive in some norm, and ② $\forall \vec{x} \in D$, $G(\vec{x}) \in D$. Then G has a unique fixed point $\vec{p}$ in G and the fixed point iteration $\vec{p}^{\,k+1} = G(\vec{p}^{\,k})$ converges to it.

$\vec{p}^{\,k} \leftarrow$ k is iteration counter

Don't confuse with coordinate of $\vec{p}$

proof

Let $\|\cdot\|$ be the norm that G is contractive in, w/ Lipschitz constant $L < 1$.

① Existence of $\vec{p}$: either G is compact, so use Brouwer, or else use "Cauchy Sequences", beyond scope of our course (take APPM 4400)

② Uniqueness: let $\vec{p} = G(\vec{p})$ and $\vec{q} = G(\vec{q})$ be fixed points.

$$\|\vec{p} - \vec{q}\| = \|G(\vec{p}) - G(\vec{q})\| \le L \cdot \|\vec{p} - \vec{q}\| < \|\vec{p} - \vec{q}\|$$

$$\Rightarrow \|\vec{p} - \vec{q}\| = 0, \text{ So } \vec{p} = \vec{q}.$$

③ (see next pg.)

* closed set: means includes its limit points (eg. boundary). In particular, all of $\mathbb{R}^n$ is closed; closed balls are closed; hyperrectangles are closed.

(proof cont'd)

Define $\vec{p} = G(\vec{p})$ the unique fixed point, $\vec{p}^{k+1} = G(\vec{p}^k)$

So $\underbrace{\|\vec{p}^{k+1} - \vec{p}\|}_{e_{k+1}} = \| G(\vec{p}^k) - G(\vec{p}) \|$

$\leq L \cdot \| \underbrace{\vec{p}^k - \vec{p}}_{e_k \text{ error}} \|$

$e_{k+1} \leq L \cdot e_k$

i.e.

(*) $e_k \leq L^k e_0$, $e_0 = \|\vec{p}^{\,0} - \vec{p}\|$ ← initial guess, and $L < 1$ so $e_k \to 0$. □

You're free to "shop around" for norms that are contractive.

<u>One tool</u>:

let $G(\vec{x}) = \begin{bmatrix} G_1(\vec{x}) \\ \vdots \\ G_n(\vec{x}) \end{bmatrix}$ $G_i : \mathbb{R}^n \to \mathbb{R}$

• if <u>each</u> G_i is L-Lipschitz in the sense $|G_i(\vec{x}) - G_i(\vec{y})| \leq L \cdot \|\vec{x} - \vec{y}\|_\infty$

then

G is L-Lipschitz w.r.t. $\|\cdot\|_\infty$ norm: $\|G(\vec{x}) - G(\vec{y})\|_\infty \leq L \cdot \|\vec{x} - \vec{y}\|_\infty$

and for each G_i, $G_i(\vec{x}) - G_i(\vec{y}) = \nabla G_i(\vec{\xi})^T (\vec{x} - \vec{y})$ (not same $\vec{\xi}$ for different i)

so $|G_i(\vec{x}) - G_i(\vec{y})| \leq \underbrace{\|\nabla G_i(\vec{\xi})\|_1 \cdot \|\vec{x} - \vec{y}\|_\infty}_{\text{via Hölder's ineq.}}$

so want $\underbrace{\|\nabla G_i(\vec{\xi})\|_1}_{= \sum_{j=1}^n \left| \frac{\partial G_i}{\partial x_j}(\vec{\xi}) \right|} \leq L < 1 \quad \forall i = 1, \ldots, n$

i.e. bound it for any input since we don't know what $\vec{\xi}$ is

(*) This error bound isn't computable since $e_0 = \|\vec{p}^{\,0} - \vec{p}\|$ isn't known.

We can fix that: $e_0 = \|\vec{p}^{\,0} - \vec{p}\| \leq \|\vec{p}^{\,0} - \vec{p}^{\,1}\| + \|\vec{p}^{\,1} - \vec{p}\|$ (triangle ineq)

$\leq \|\vec{p}^{\,0} - \vec{p}^{\,1}\| + L \cdot \underbrace{\|\vec{p}^{\,1} - \vec{p}\|}_{e_0}$

so $(1 - L) e_0 \leq \|\vec{p}^{\,0} - \vec{p}^{\,1}\|$

i.e. $e_0 \leq \frac{1}{1-L} \underbrace{\|\vec{p}^{\,0} - \vec{p}^{\,1}\|}_{\text{computable}}$