Class #11

Friday, February 28, 2025 9:34 AM

Newton Decrement: $\lambda(x) = \left[\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)\right]^{1/2}$

When we are close to a minimizer p^*

$\frac{\lambda(x)^2}{2} \Rightarrow$ Estimate of how close we are to p^*

$$f(x) - p^* \approx \frac{\lambda(x)^2}{2}$$

Some facts: ① $\nabla f(x)^T \underline{\Delta x_{nt}} = -\lambda(x)^2$

$\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x)$

$$\|\nabla f(x)\|_2 \|\Delta x_{nt}\|_2 \cos(\theta)$$

② $\lambda(x) = \left(\Delta x_{nt} \nabla^2 f(x) \Delta x_{nt}\right)^{1/2}$

P-Quadratic norm of a vector when $P \in S^n_{++}$

$$\|x\|_P = \left(x^T P x\right)^{1/2}$$

$$\lambda(x) = \|\Delta x_{nt}\|_{\nabla^2 f(x)}$$

③ $\lambda(x)$ is invariant under affine transformations of the function $f(x)$

Newton's step:

$$x^{(k+1)} = x^{(k)} + t^{(k)} \underbrace{\Delta x_{nt}}_{-\underbrace{\nabla^2 f(x)^{-1}}_{\nabla^2 f(x) \succ 0} \nabla f(x)}$$

when $t^{(k)} = 1 \Rightarrow$ Full Newton Step.

↳ Full Newton Step should be taken only when x is very close to the local (or global) minimum. Otherwise, the method can diverge.

Newton's Method (Assuming we are close to a local minimum or are working with a convex problem)

Input: $x^{(0)} \in \text{dom} f$; tolerance $\epsilon > 0$

Initialize: $k \leftarrow 0$

Repeat:

① $\Delta x_{nt}^{(k)} \leftarrow -\underbrace{\nabla^2 f(x^{(k)})^{-1}}_{} \nabla f(x^{(k)})$ // Compute Newton's direction

↳ make sure $\nabla^2 f(x^{(k)}) \succ 0$

② Compute $\lambda^2(x^{(k)}) \leftarrow \nabla f(x^{(k)})^T \nabla^2 f(x^{(k)})^{-1} \nabla f(x^{(k)})$

// Compute Newton decrement Square

③ If $\frac{\lambda^2(x^{(k)})}{2} \leq \epsilon$

break

④ <u>Line Search</u>: Choose $t^{(k)}$ using back tracking line search

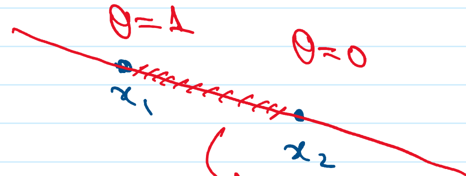
⑤ <u>Update</u>: $x^{(k+1)} \leftarrow x^{(k)} + t^{(k)} \Delta x_{nt}^{(k)}$

⑥ $k \leftarrow k+1$

~~~~~~~~ × ~~~~~~~~ × ~~~~

<u>Chapter 2</u>: Convex Sets

① <u>Line</u>: Given $x_1 \neq x_2 \in \mathbb{R}^n$



$$y = \theta x_1 + (1-\theta) x_2 \; ; \; \theta \in \mathbb{R}$$

$\theta \in [0,1]$

$y(\theta)$

↳ one-dimensional linear function

$y(\theta)$ is a line passing through x_1 and x_2

② <u>Line Segment</u>: when $\theta \in [0,1]$,

$y(\theta) = \theta x_1 + (1-\theta) x_2$; is the line segment joining x_1 and x_2
~~~~~~~~

$$\downarrow$$

$$= x_2 + \theta(x_1 - x_2)$$

$\theta \Rightarrow$ Tells us how far we are from x_2 in the direction $(x_1 - x_2)$

③ <u>Affine Set</u>: A set $C \subseteq \mathbb{R}^n$ is affine if the line through any distinct points in C lies in the set C.

$$\forall\ x_1, x_2 \in C \text{ and } \theta \in \mathbb{R}$$

$$\theta x_1 + (1-\theta)x_2 \in C$$

↳ Linear combinations of points in C lie in C provided the coefficients of the linear combination sum to 1.

$\Downarrow$ simple induction argument

If C is an affine set, $x_1, \ldots, x_k \in C$, $\theta_1, \ldots, \theta_k \in \mathbb{R}$ **s.t.** $\sum_{j=1}^{k} \theta_j = 1$ then

$$\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \in C$$

<u>Affine Combination of Points</u>

Given $x_1, \ldots, x_k \in \mathbb{R}^n$ and $\theta_1, \ldots, \theta_k \in \mathbb{R}$

s.t. $\sum_{j=1}^{k} \theta_j = 1$

$\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$ is called an affine combination.

Connection to a Subspace

An affine set is a subspace shifted from the origin.

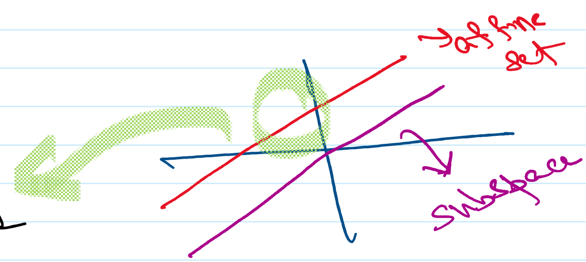
If C is an affine set and $x_0 \in C$ for any x_0 then

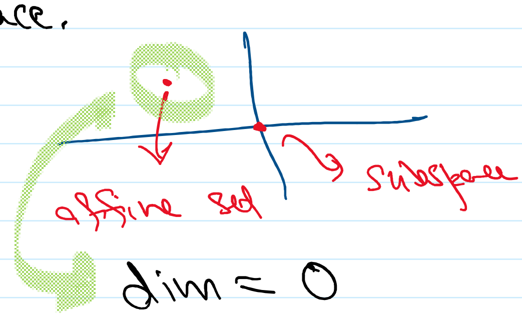
$$V = C - x_0 = \{x - x_0 : x \in C\}$$

↳ Then V is a subspace.

$\mathbb{R}^2$

dim = 1





dim = 0

Dimension of an Affine Set

It is the dimension of the subspace associated with C.

Ex: $C = \{x : Ax = b ; b \in \mathbb{R}^n\}$

Is C affine?

Approach 1: Use the definition of an affine set.

Let x_1 and $x_2 \in C$

Show that $\theta x_1 + (1-\theta) x_2 \in C$

$$x_1 \in C \Rightarrow Ax_1 = b$$

$$x_2 \in C \Rightarrow Ax_2 = b$$

$$\theta x_1 + (1-\theta) x_2 \Rightarrow \in C \,?$$

Show $\Rightarrow A(\theta x_1 + (1-\theta) x_2) = b$

$$= \theta \underbrace{Ax_1}_{b} + (1-\theta) \underbrace{Ax_2}_{b}$$

$$= \theta b + (1-\theta) b = b \checkmark$$

<u>Approach 2:</u> $\underbrace{\{x : Ax = 0\}}_{\text{null space of } A}$ is a subspace

$\{x : Ax = b\}$ is simply a shifted version of $\{x : Ax = 0\} \Rightarrow C$ is an affine set.

Dimensionality of C = dimensionality of the null space of A.

<u>Affine Hull of points in a set in $\mathbb{R}^n$</u>

Given a set C, which may not be affine, the affine hull of C is defined as:

Given a set C, which may not be affine, the affine hull of C is defined as:

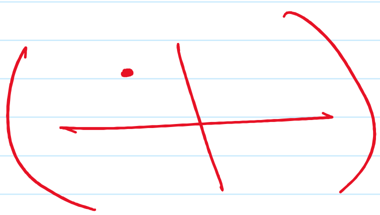
$$\operatorname{aff} C = \{ \theta_1 x_1 + \dots + \theta_k x_k : x_1, \dots, x_k \in C, \ \theta_1 + \theta_2 + \dots + \theta_k = 1 \}$$

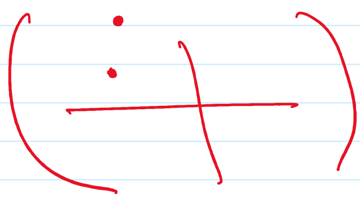
↳ Affine hull of a set C is the smallest affine set that contains C. That is, let S be any affine set such that:

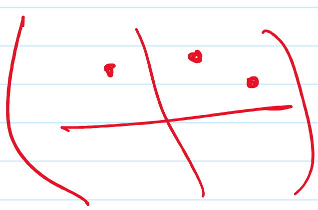
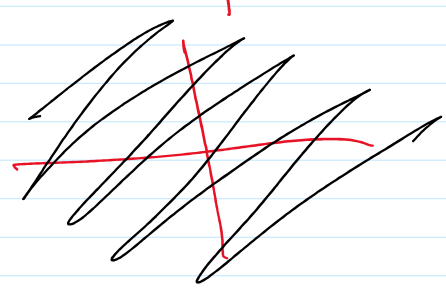
$$C \subseteq S. \text{ Then,}$$

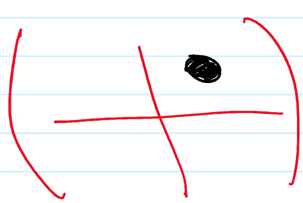
$$\operatorname{aff} C \subseteq S$$

Examples:

aff  = 

aff  = 

aff  = 

aff  $= \mathbb{R}^2$

Reading: 2.1.1 and 2.1.2 (BV)

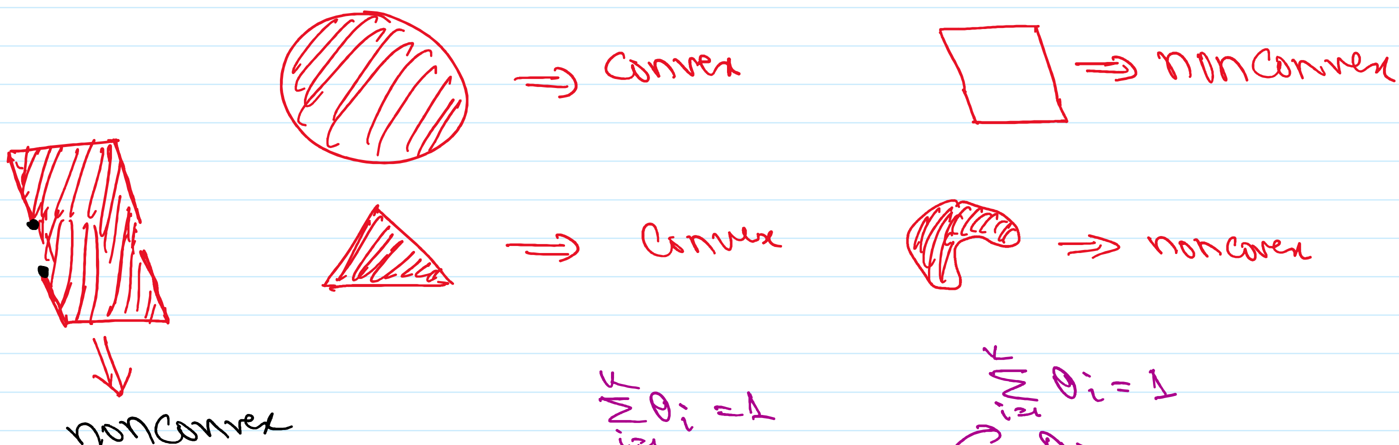
Convex Set

A set C is convex if the line segment between any two points in C, lies in C.

$$\forall\ x_1, x_2 \in C \text{ and } \theta \in [0,1]$$

$$\theta x_1 + (1-\theta) x_2 \in C$$

- Every affine set is a convex set



linear combination

$\sum_{i=1}^{k} \theta_i = 1$

$\sum_{i=1}^{k} \theta_i = 1$, $\theta_i \geq 0$

Subspaces $\subset$ affine sets $\subset$ Convex sets

Convex Combination of Points

Given $x_1, x_2, \dots, x_k \in C$ and $\theta_1, \dots, \theta_k \geq 0$

s.t. $\sum_{j=1}^{k} \theta_j = 1$

$\theta_1 x_1 + \cdots + \theta_k x_k$ is called a convex combination of the points in C.