Gradients, Jacobians, Hessians

Gradient: Given $f: \mathbb{R}^n \to \mathbb{R}$

the gradient is $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$

Example:

$f(x) = a^T x, \quad a \in \mathbb{R}^n, x \in \mathbb{R}^n$

$= a_1 x_1 + \dots + a_n x_n.$

We can compute the gradient

$$\nabla f(x) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a$$

If our function has multiple outputs, $g: \mathbb{R}^n \to \mathbb{R}^m$, we can define the Jacobian.

We have $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix}$

The Jacobian is an $m \times n$ matrix, given by

$$Dg = \begin{pmatrix} \nabla g_1 \\ \vdots \\ \nabla g_m \end{pmatrix}$$

Example $g(x) = Ax, \quad A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n$

We want to compute Dg

rows of A

We have

$$g(x) = Ax = \begin{pmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{pmatrix} x = \begin{pmatrix} a_1^T x \\ \vdots \\ a_m^T x \end{pmatrix}$$

To compute Jacobian, we have

$$Dg(x) = \begin{pmatrix} \nabla(a_1^T x) \\ \vdots \\ \nabla(a_m^T x) \end{pmatrix} = \begin{pmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{pmatrix} = A$$

Now if we have

$$h(x) = x^T A x,$$

to compute gradient of h, we can use product rule:

We have $D(f(x)^T g(x))$

$$= f(x)^T Dg(x) + g(x)^T Df(x)$$

In this example, $f(x) = x$, $g(x) = Ax$

$Df(x) = I$ $\quad Dg(x) = A$

So

$$D(x^T A x) = x^T A + \cancel{I} \cdot x^T A^T$$

$$= x^T (A + A^T)$$

And

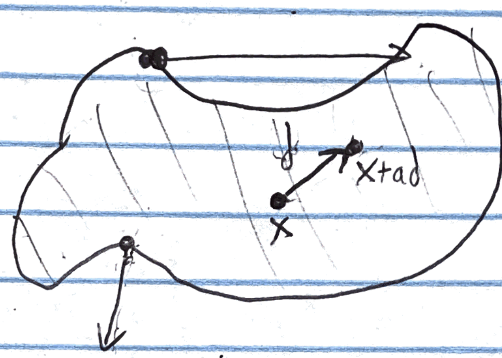
$$\nabla(x^T A x) = (A + A^T) x$$

If A is symmetric, $A = A^T$,

and $\nabla(x^T A x) = 2Ax$

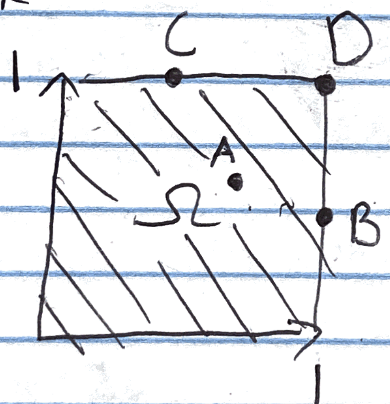
Feasible Directions:

At a point $x \in \Omega$, a vector $d \neq 0 \in \mathbb{R}^n$ is <u>feasible</u> if $\exists\, \alpha > 0$ s.t. for all $a \leq \alpha$, $x + ad \in \Omega$

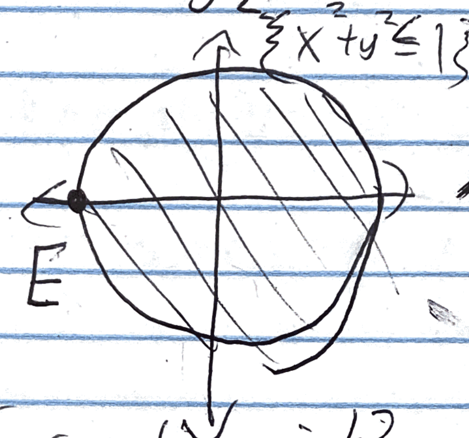


Partner work

A point is interior if all directions are feasible.



$\Omega = \{x^2 + y^2 \leq 1\}$



What are the feasible directions for each point?

A: $\mathbb{R}^2 \setminus \{0\}$

B: $\{(x,y): x \leq 0\} \setminus \{0\}$

C: $\{(x,y): y \leq 0\} \setminus \{0\}$

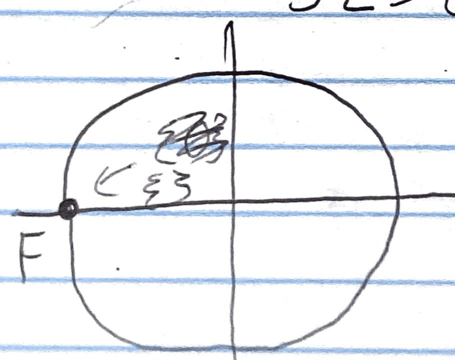
D: $\{(x,y): x, y \leq 0\} \setminus \{0\}$

E: $\{(x,y): x > 0\}$

F: $\{\}$

What about

$\Omega = \{x, y: x^2 + y^2 =$

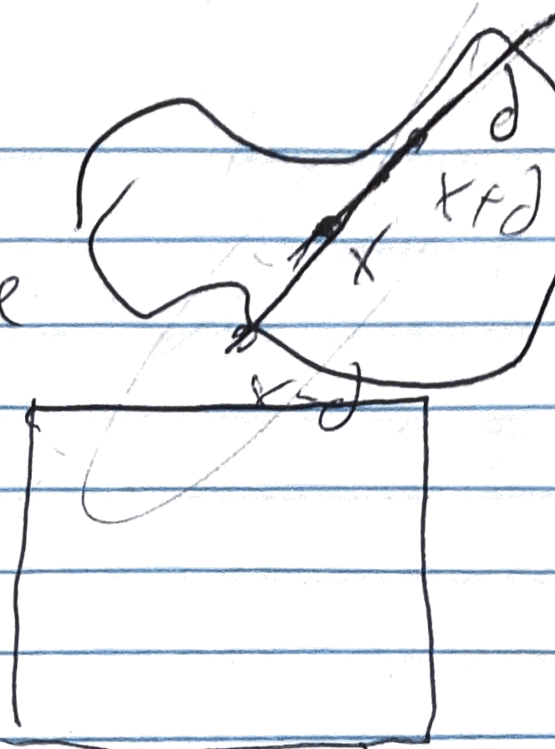


~~Basic Strategy To~~

Review: Directional Derivative

~~What~~

Given direction d, what is derivative in direction d at point x?



$g(a) = f(x+ad)$ $\quad g'(a) = \nabla f(x+ad) \cdot d$

$g'(0) = \nabla f(x) \cdot d$

↑ Maximized when $d = \nabla f(x)$,

Hence $\nabla f(x)$ is direction of maximum increase.

FONC (Interior): $\nabla f(x) = 0$

FONC (Boundary): $d^T \nabla f(x) \geq 0$ for all feasible d

"All feasible directions are increasing"