(Unconstrained) Convex Optimization

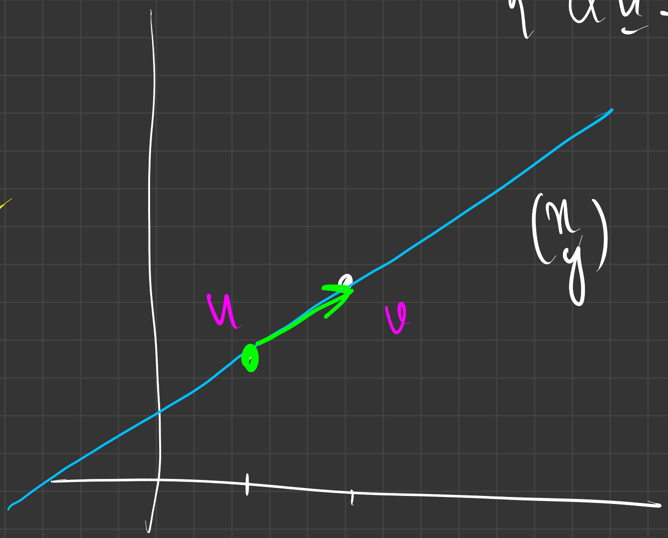
Line

$\underline{u}, \underline{v} \in \mathbb{R}^n$

The straight line passing through $\underline{u}$ & $\underline{v}$

$$= \{ \alpha \underline{u} + (1-\alpha)\underline{v} : \alpha \in \mathbb{R} \}$$

$$f(\alpha) = \underline{u} + \alpha \underline{v}$$



$$\underline{x} = \underline{u} + (1-\alpha)(\underline{v}-\underline{u})$$

$$= \alpha \underline{u} + (1-\alpha)\underline{v}$$

$(x_1, y_1) \quad (x_2, y_2)$

$$\frac{(y-y_2)}{(x-x_2)} = \frac{(y_1-y_2)(1-\alpha)}{(x_1-x_2)(1-\alpha)}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + (1-\alpha)\begin{pmatrix} x_1-x_2 \\ y_1-y_2 \end{pmatrix}$$

$(1, 2) \quad (2, 3)$

$\downarrow \quad \downarrow$

$\underline{u} \quad v$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (1-\alpha)\begin{pmatrix} 2-1 \\ 3-2 \end{pmatrix}$$

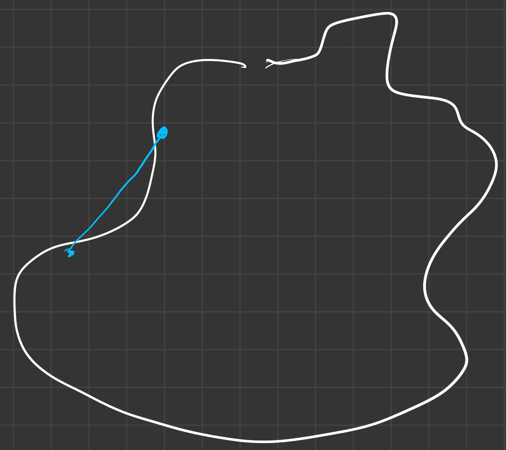
$$= \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (1-\alpha)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

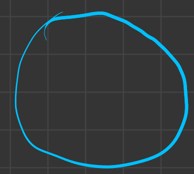
$=$

$\{ \alpha \underline{u} + (1-\alpha)\underline{v} : \alpha \in [0,1] \} \longrightarrow$ line segment joining $\underline{u}$ & $\underline{v}$

Convex set: A is convex if $\underline{u}, \underline{v} \in A$

$\rightarrow \alpha \underline{u} + (1-\alpha)\underline{v} \in A \quad \forall \alpha \in [0,1]$





Convex functions

If $f: \mathbb{R}^n \to \mathbb{R}$

$f: S \to \mathbb{R}$ $\quad S \subseteq \mathbb{R}^n$

(S is a convex set)

f is convex if

$$f(\alpha \underline{x} + (1-\alpha)\underline{y}) \leq \alpha f(\underline{x}) + (1-\alpha) f(\underline{y})$$

$\forall \alpha \in [0,1]$

$\underline{x}, \underline{y} \in S$

Eg: $\quad f(\underline{x}) = \|x\|_2^2 = \sum_{i=1}^{n} x_i^2$

<u>Claim</u>: f_1 & f_2 are convex. Then, $f_1 + f_2$ is convex

$$f(x) = f_1(x) + f_2(x)$$

$$f(\alpha x_1 + (1-\alpha)x_2) = f_1(\alpha x_1 + (1-\alpha)x_2) + f_2(\alpha x_1 + (1-\alpha)x_2)$$

$$\leq \alpha f_1(x_1) + (1-\alpha) f_1(x_2) + \alpha f_2(x_1) + (1-\alpha) f_2(x_2)$$

$$= \alpha(f_1(x_1) + f_2(x_1)) + (1-\alpha)(f_1(x_2) + f_2(x_2))$$

$$= \alpha f(x_1) + (1-\alpha) f(x_2)$$

Unconstrained minimization of a convex function

$f : \mathbb{R}^n \to \mathbb{R}$ convex, differentiable

$$x^* = \underset{x \in \mathbb{R}^n}{\arg\min} \; f(x)$$

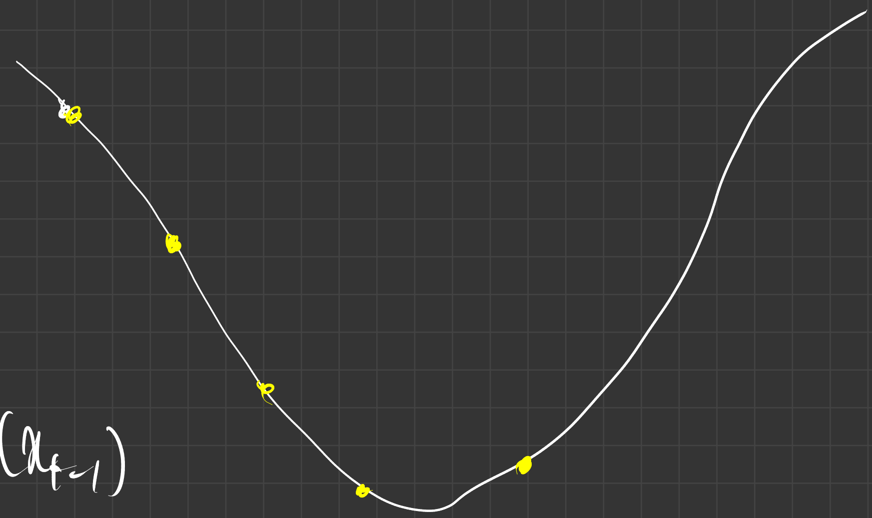
$$\to \quad \nabla f(x^*) = 0$$

Gradient descent

$\underline{x}_0 \in \mathbb{R}^n$

for $t = 1, 2, \ldots$:

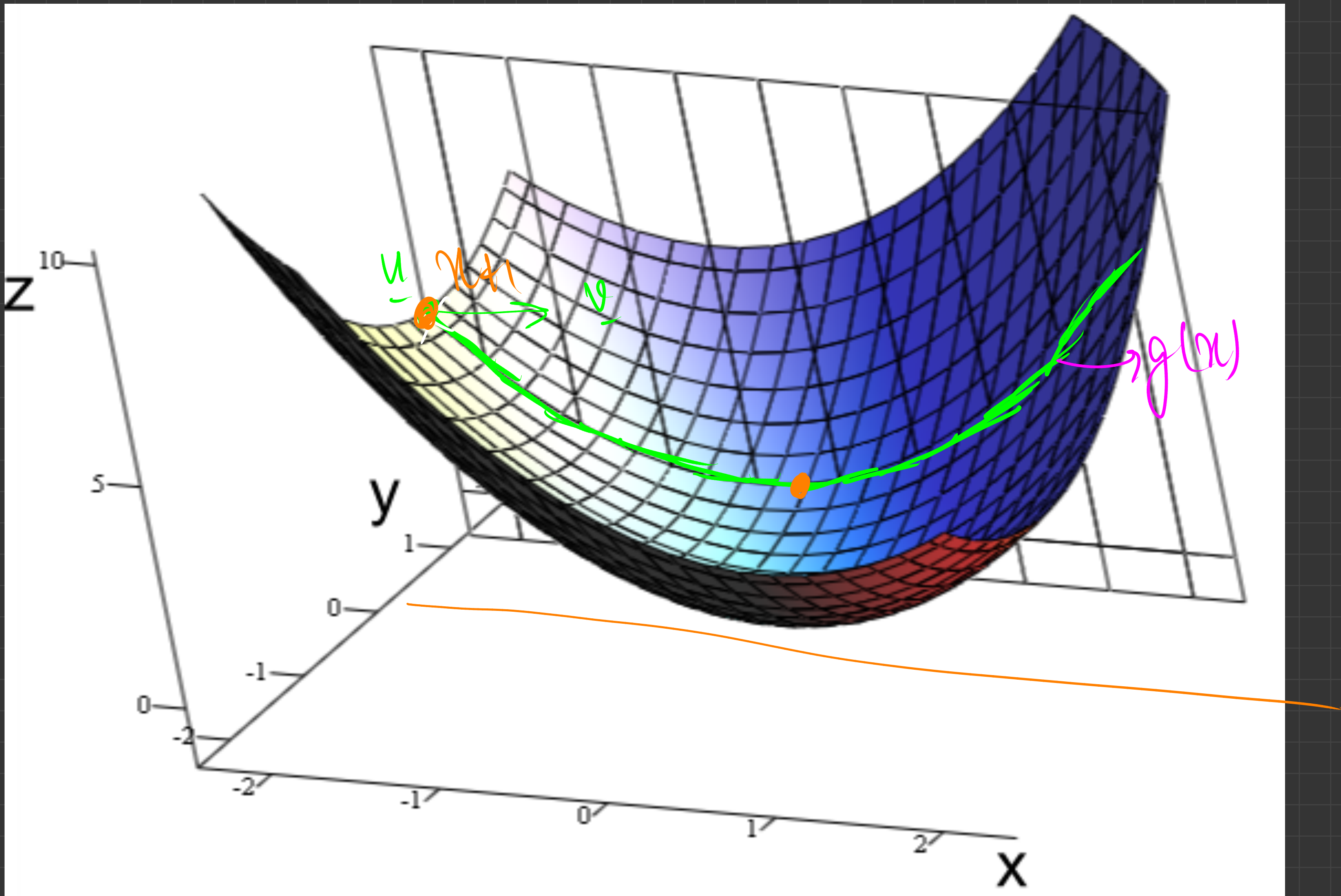
$$\underline{x}_t = \underline{x}_{t-1} - \gamma_t \nabla f(\underline{x}_{t-1})$$



Claim: The direction where the tangent has min slope $= -\nabla f(\underline{x})$

$$f(\underline{x}_t) = f(\underline{x}_{t-1} - \gamma_t \nabla f(\underline{x}_{t-1}))$$

Slices of convex functions are convex



Claim: $f : \mathbb{R}^n \to \mathbb{R}$

$$g(x) = f(\underline{u} + x\underline{v}) \quad \text{for fixed } \underline{u}, \underline{v}$$

$$g : \mathbb{R} \to \mathbb{R}, \quad x \in \mathbb{R}$$

$$x_1, x_2 \in \mathbb{R} \ \& \ \alpha \in [0,1]$$

$$g(\alpha x_1 + (1-\alpha)x_2) = f(\underline{u} + (\alpha x_1 + (1-\alpha)x_2)\underline{v})$$

$$= f((\alpha + (1-\alpha))\underline{u} + (\alpha x_1 + (1-\alpha)x_2)\underline{v})$$

$$= f(\alpha \underbrace{(\underline{u} + x_1\underline{v})}_{y} + (1-\alpha)\underbrace{(\underline{u} + x_2\underline{v})}_{z})$$

$\le$ (By convexity of f)

$$\le \alpha f(\underline{u} + x_1\underline{v}) + (1-\alpha) f(\underline{u} + x_2\underline{v})$$

$$= \alpha g(x_1) + (1-\alpha) g(x_2)$$

Method of steepest descent

$\underline{x}_0$

for $t = 1, 2, 3 \ldots$

$$\alpha_t = \operatorname*{argmin}_{\alpha \geq 0} f(\underline{x}_{t-1} - \alpha \nabla f(\underline{x}_{t-1}))$$

$$\underline{x}_t = \underline{x}_{t-1} - \alpha_t \nabla f(\underline{x}_{t-1})$$


x_0
x_1

Properties

Prop 1: Say $\underline{x}_0, \underline{x}_1, \underline{x}_2 \ldots$ are points obtained by the steepest descent algorithm. Then, for any $t = 1, 2, \ldots$

$$(\underline{x}_t - \underline{x}_{t+1})^T (\underline{x}_{t+1} - \underline{x}_t) = 0$$

Proof:

$$\alpha_{t+1} = \operatorname*{argmin}_{\alpha} f(\underline{x}_t - \alpha \nabla f(\underline{x}_t))$$

$$① \quad \frac{d}{d\alpha} f(\underline{x}_t - \alpha \nabla f(\underline{x}_t)) \Big|_{\alpha_{t+1}} = 0$$

$$(\underline{x}_t - \underline{x}_{t-1})^T (\underline{x}_{t+1} - \underline{x}_t) = (\alpha_t \nabla f(x_{t-1}))^T (\alpha_{t+1} \nabla f(x_t))$$

$$= \alpha_t \alpha_{t+1} (\nabla f(x_{t-1}))^T (\nabla f(x_t))$$

$$\frac{d}{d\alpha} \overbrace{f(\underbrace{\underline{x}_t - \alpha \nabla f(\underline{x}_t)})}^{g(\alpha)} \Big|_{\alpha_{t+1}} = 0$$

$$g(\alpha) = g_1(g_2(\alpha)) \qquad g_2(\alpha) = \underline{x}_t - \alpha \nabla f(x_t)$$

$$g_1(\underline{x}) = f(\underline{x})$$

$$g_2 : \mathbb{R} \to \mathbb{R}^n$$

$$g_1 : \mathbb{R}^n \to \mathbb{R}$$

$$0 = \frac{dg(\alpha)}{d\alpha}\Big|_{\alpha=\alpha_{t+1}} = Dg_1(\underline{x})\, Dg_2(\alpha)\Big|_{\alpha_{t+1}}$$

$$= \left(\nabla f(\underline{x})\right)^T \left(-\nabla f(\underline{x}_t)\right)\Big|_{\alpha_{t+1}}$$

$$\underline{x} = \underline{x}_t - \alpha \nabla f(\underline{x}_t)$$

$$= -\left(\nabla f(\underline{x}_{t+1})\right)^T \left(\nabla f(\underline{x}_t)\right)$$

Property 2: Suppose f is convex & $\nabla f(x) \neq 0 \;\forall\, \underline{x} \neq \underline{x}_{min}$

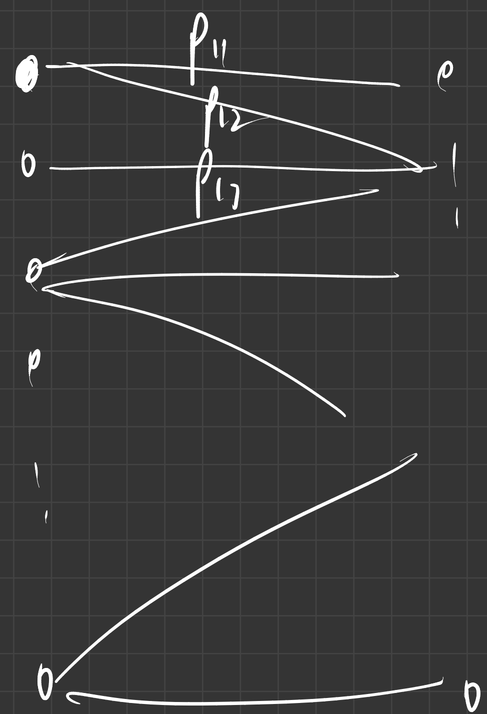
Then, if $x_0, x_1, \ldots$ are points in the M.O.S.D,

$$f(x_t) < f(x_{t-1})$$

$\Rightarrow$ f is also bounded, then MOSD will converge in a finite # of steps.

$$\min_{x \in C} f(x)$$

Bipartite graph matching



n students n faculty

p_{ij} : pref of i for j

Goal: Ensure that each student is matched with 1 fac

each fac — " — 1 student

$$\max \sum_{\substack{i \& j \\ \text{matched}}} p_{ij}$$