Math Foundations for ML

10-606

Geoff Gordon

Notes and reminders

- HW2 is out

$x \in \mathbb{R}^n \quad f \in \mathbb{R}^n \to \mathbb{R}^n$

Nonlinear systems

$f'(x_1) \in \mathbb{R}^{n \times n}$

start w/ x_1

$$df(x) = f'(x)\,dx$$

$$f(x_1) + f'(x_1)\,dx = 0$$

→ solve for dx

$$x_2 := x_1 + dx$$

$$f(x_2) + f'(x_2)\,\underline{dx} = 0$$

$\vdots$

$$f(x) = e^x - 1$$

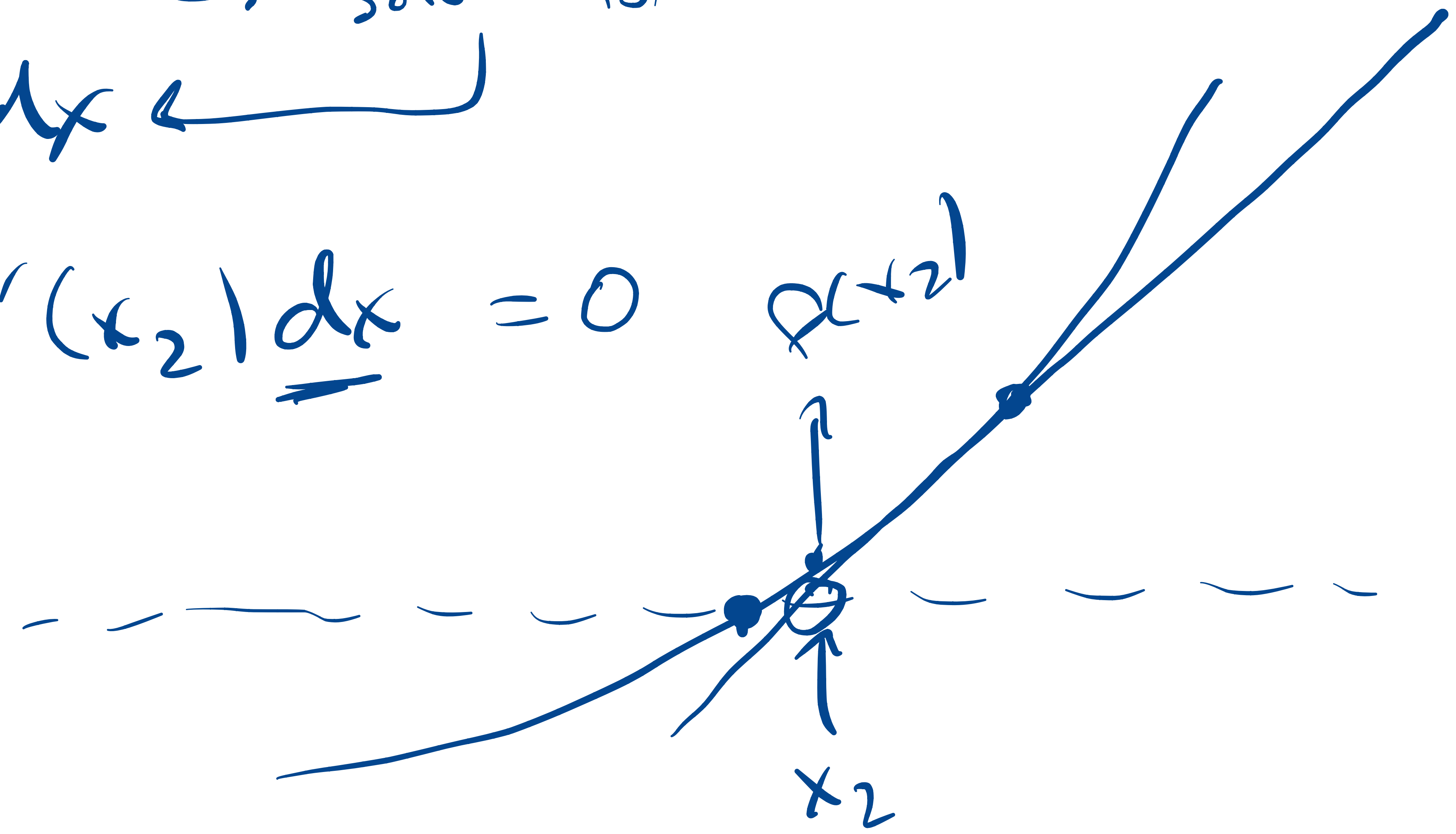
solve $f(x) = 0$

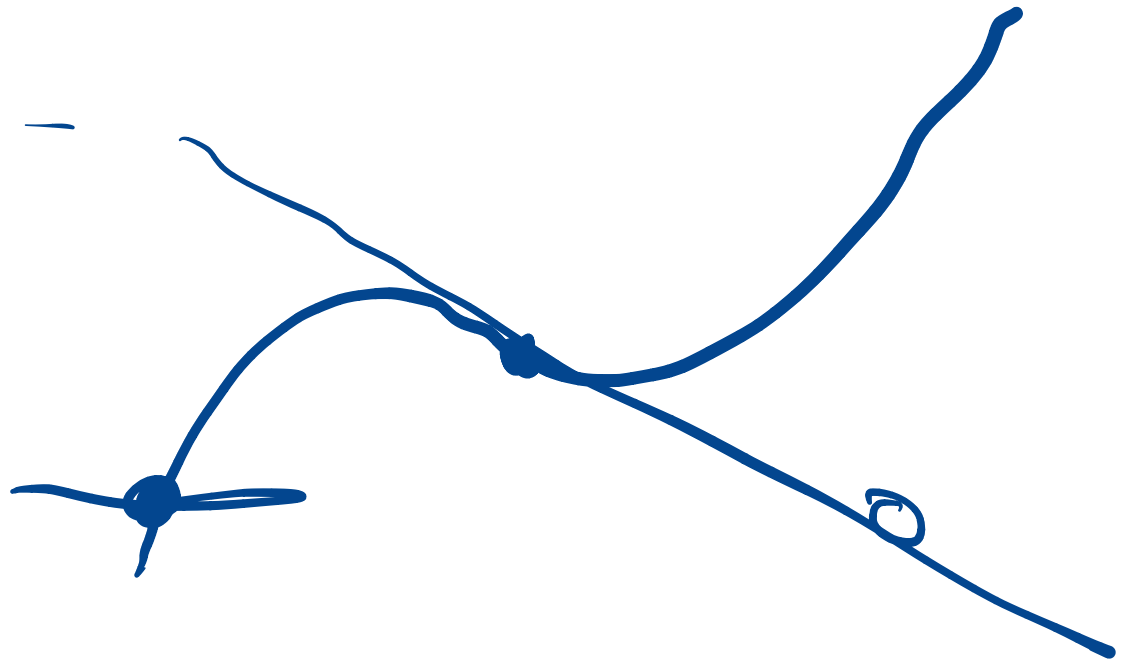
$$df(x) = e^x\,dx$$

What is $f(x+dx)$?

$$df = f'(x)\,dx$$

$$\underline{f + df} = f + f'(x)\,dx$$





Newton's method

f(x) = e^x–1

x	f	df	Equation	dx
1	$e-1$	e	$e\,dx = 1-e$	$\frac{1-e}{e}$
-0.632	-0.468	0.532	$0.532\,dx = 0.468$	0.880
0.248	0.281	1.281	$1.281\,dx = -0.281$	-0.219

final: x=0.029

Exercise

on repl.it

Mc Cullagh & Nelder

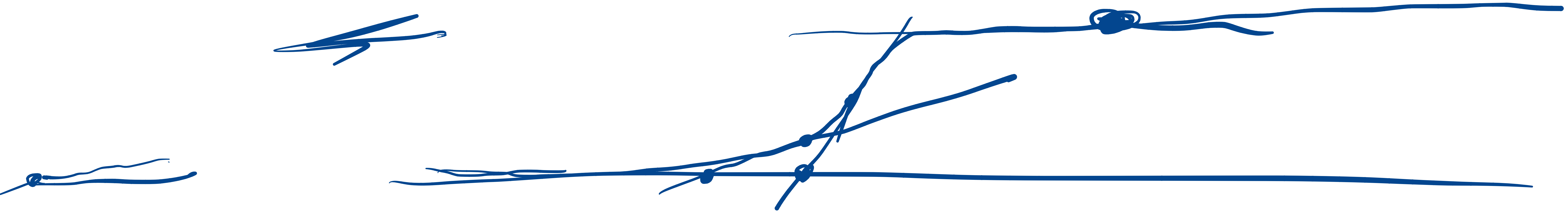
IRLS

↳ Newton for logistic

$f_i + df = 0$

$df = f'(x_i)\,dx$

- Try out Newton's method for one of the functions provided
- Then pick a different simple function and implement your own version of newton_step; see what happens for various initial points
- https://replit.com/team/professorgeoff/Newtons-method



Unconstrained optimization

$$\min_{\theta} L(\theta)$$

$$L(\theta) \in \mathbb{R}^n \to \mathbb{R}$$

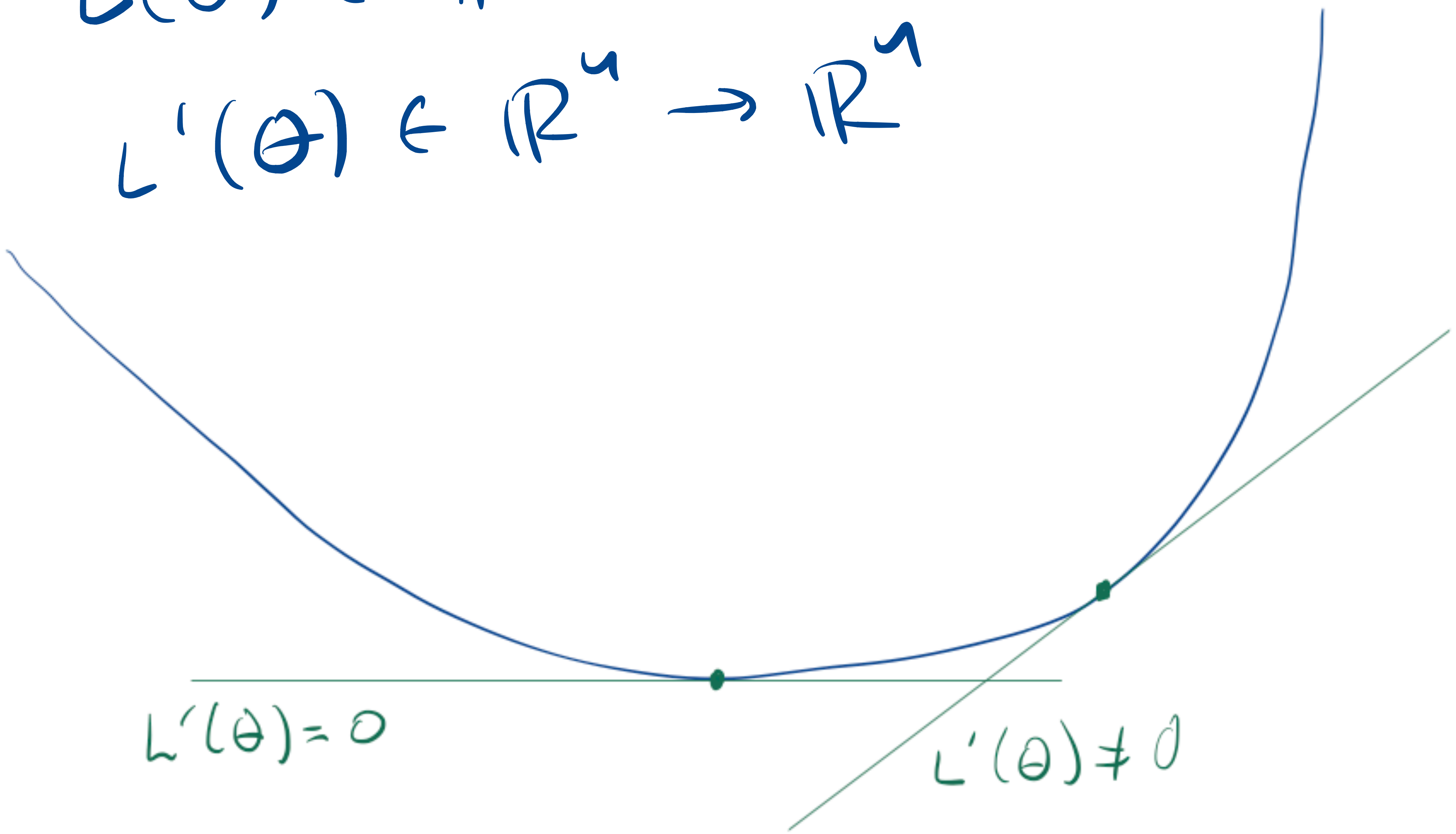
$$L'(\theta) \in \mathbb{R}^n \to \mathbb{R}^n$$

$$dL = L'(\theta)\, d\theta$$

$$L'(\theta) = 0$$

$$dL'(\theta) = L''(\theta)\, d\theta$$

$$L'(\theta_i) + \underbrace{dL'}_{L''(\theta_i)\, d\theta} = 0$$



$$\min_{x} e^x - x \qquad e^x - 1 = 0$$

Constrained optimization

$\min_\theta L(\theta)$ s.t. $g(\theta) = 0$

$L: \mathbb{R}^n \to \mathbb{R}$ $g: \mathbb{R}^n \to \mathbb{R}$ or $\to \mathbb{R}^d$

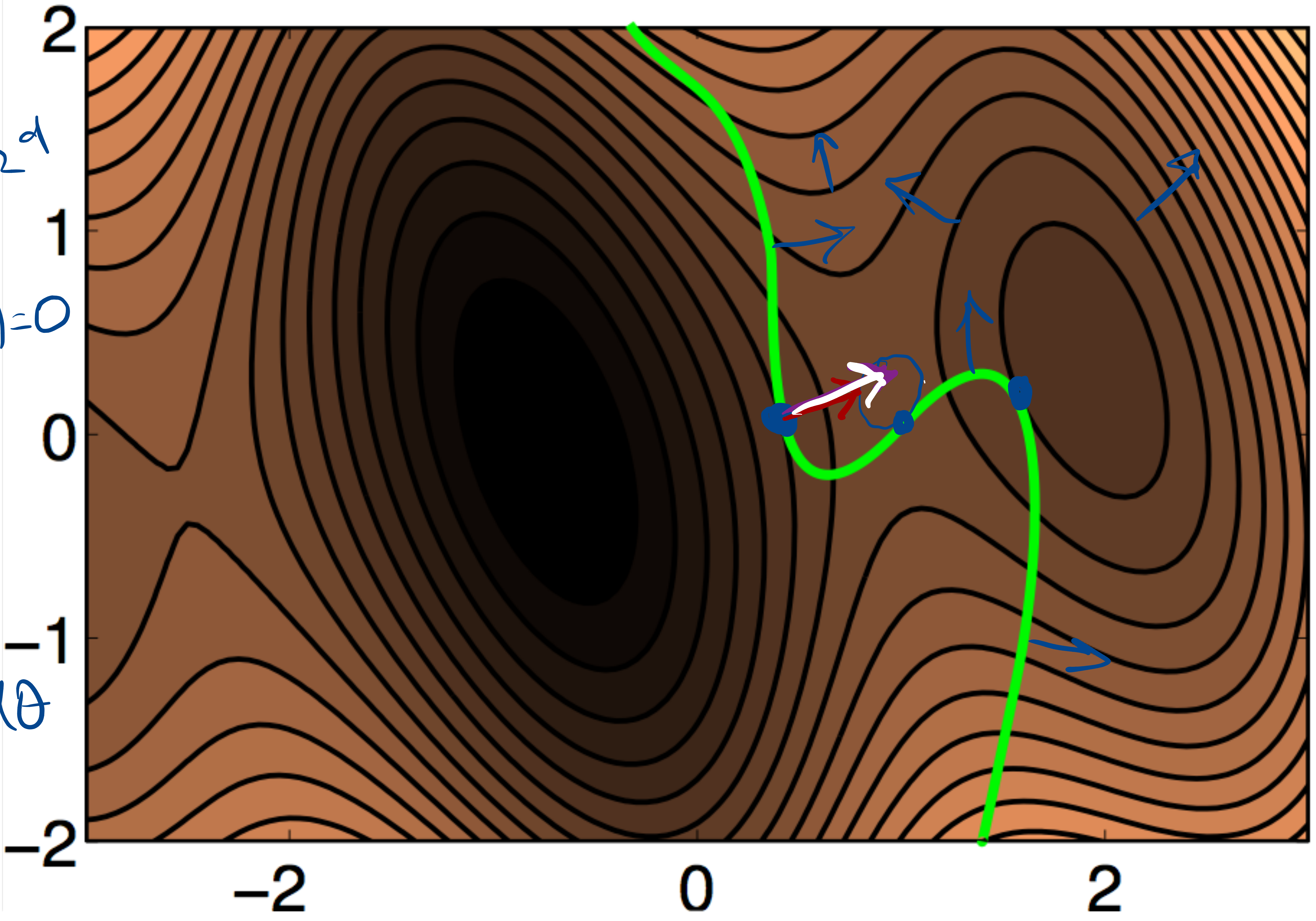
$\min_\theta L(\theta) + \alpha g(\theta)$ s.t. $g(\theta) = 0$

$d(L(\theta) + \alpha g(\theta))$

$= dL(\theta) + \alpha\, dg(\theta)$

$= L'(\theta)\, d\theta + \alpha g'(\theta)\, d\theta$

$\equiv 0$



$$\left.\begin{aligned} L'(\theta) + \alpha g'(\theta) &= 0 \\ g(\theta) &= 0 \end{aligned}\right]$$ first order optimality

$g(\theta)\ \mathbb{R}^q \to \mathbb{R}^d$

$\alpha \in \mathbb{R}^{1\times q}$

$$\min \underbrace{\tfrac{1}{2}(x^2 + y^2 + z^2)}_{L} \quad \text{st} \quad \overbrace{x+y}^{g_1} = 1, \quad \underbrace{y+z}_{g_2} = 1$$

$$dL = x\,dx + y\,dy + z\,dz = (x\ y\ z)\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$$

$$L' = (x, y, z)$$

$$g_1(x,y,z) = x+y-1$$

$$dg_1 = (1\ 1\ 0)\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$$

$$dg_2 = (0\ 1\ 1)\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix}$$

$$\left.\begin{aligned} &(x, y, z) + \alpha(1\ 1\ 0) + \beta(0\ 1\ 1) = 0 \\ &x+y = 1 \\ &y+z = 1 \end{aligned}\right]$$

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

$x = 1/3$

$y = 2/3$

$z = 1/3$

$\alpha = -1/3$

$\beta = -1/3$