Math Foundations for ML

10-606

Geoff Gordon

Reminders

{'name': 'value', 'hi': 'thee'}

- HW1 due Wed
    - written part through Gradescope, programming through repl.it
    - should be able to access these even if you aren't registered; contact us for invites if necessary
- Quiz1 on Fri in class
    - we'll save some time for review Wed; bring questions and post on Piazza
    - 80 min, all written problems, closed book/notes

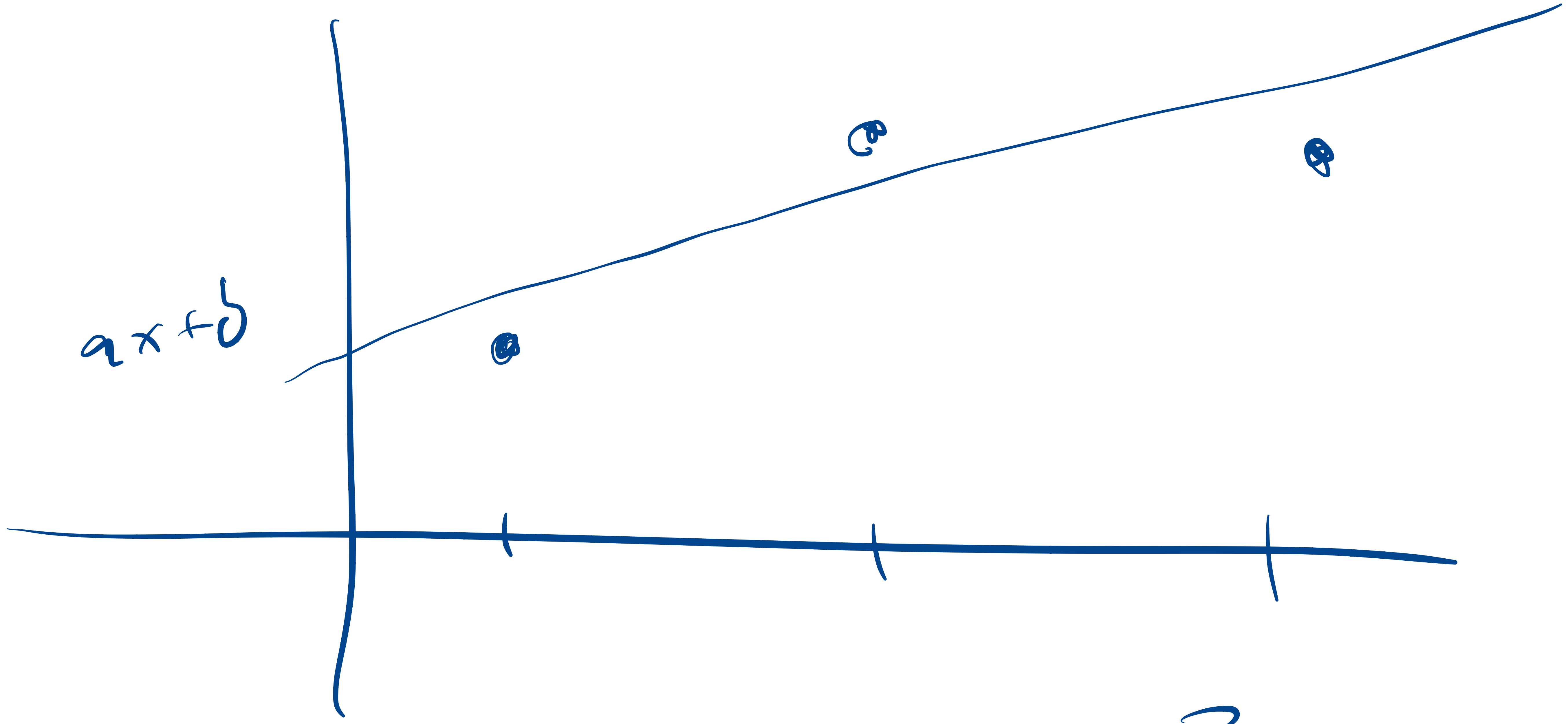
$$L(\underset{\sim}{w}) = \sum_{t=1}^{3} \left(y_t - (a x_t + b)\right)^2$$

$\hookrightarrow (a)$

$L(w, b)$

$L(v)$

$\hookrightarrow v = \begin{pmatrix} w \\ b \end{pmatrix}$



$$R(w) = \lambda \|w\|^2$$

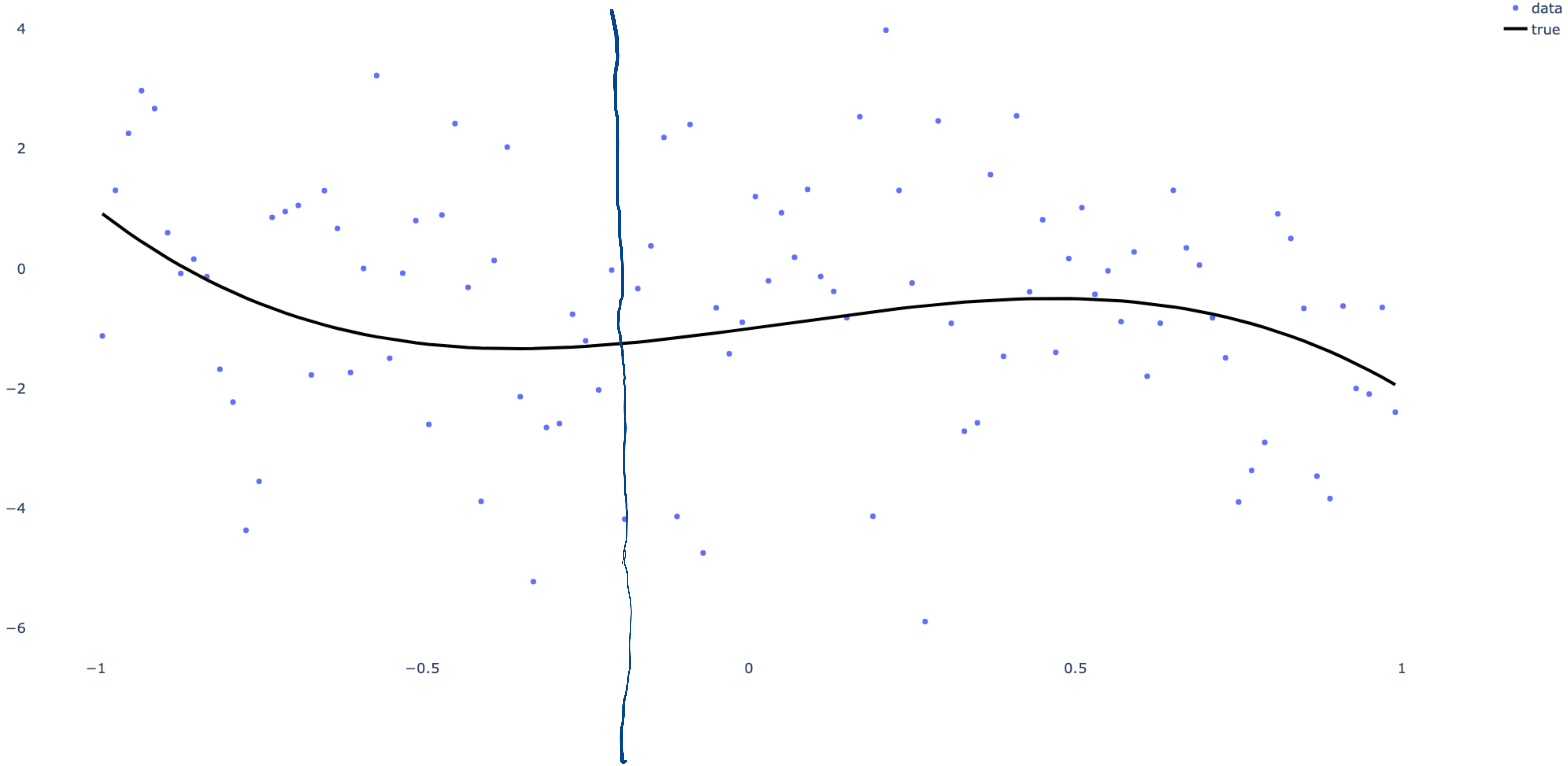
$$x \rightarrow \phi(x)$$

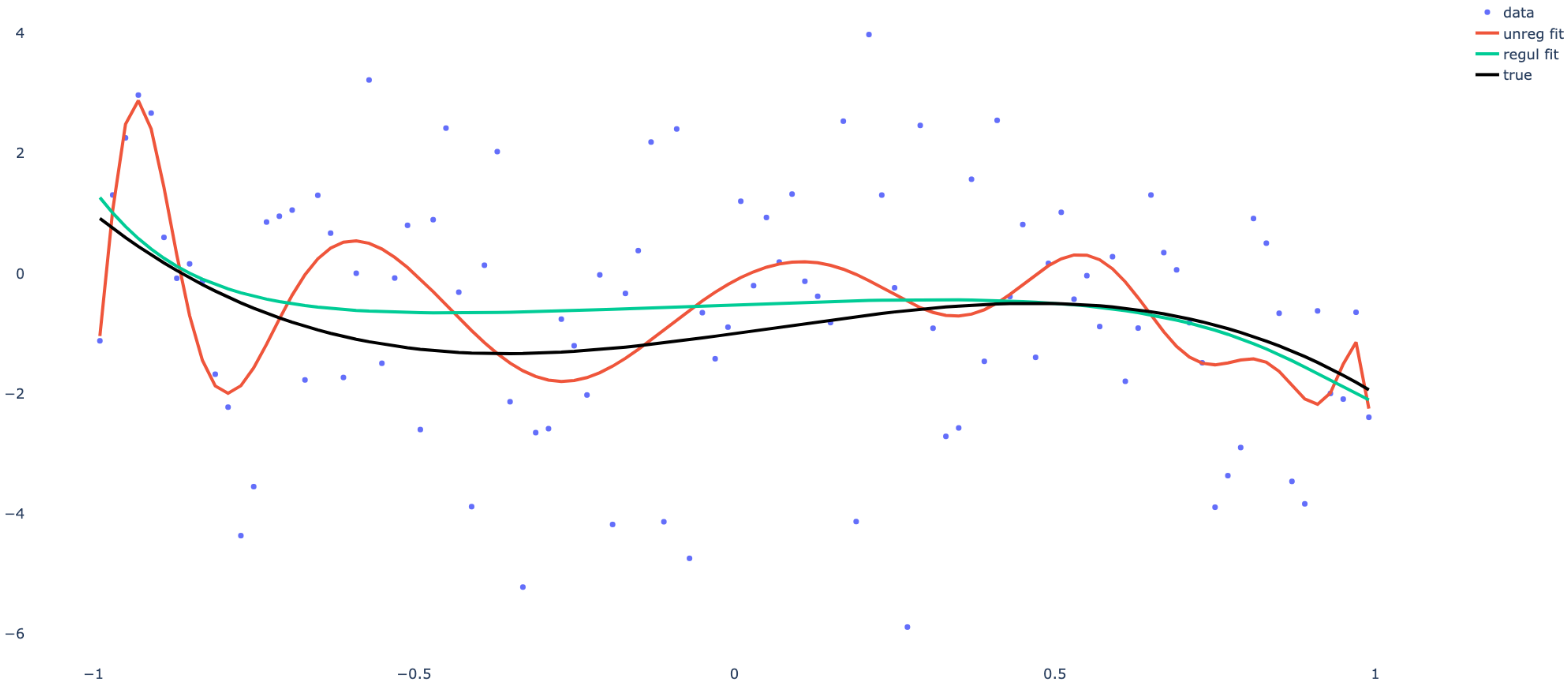
$$\begin{pmatrix} x^2 \\ x^3 \\ x \\ 1 \end{pmatrix}$$

$$w^T (X X^T + \lambda I) = y X^T$$

$\hookrightarrow$ $\left(\; | \; | \; | \; | \; x_t \; | \; | \; | \; \right)$

Linear regression





4
2
0
−2
−4
−6
−1
−0.5
0
0.5
1
data
unreg fit
regul fit
true

$$w X X^T = y X^T \qquad X = U \Sigma V^T$$

$$w U \Sigma V^T V \Sigma U^T = y V \Sigma U^T$$

($V^T V = I$)

$$w U \Sigma^2 U^T = y V \Sigma U^T$$

$$\boxed{w U \Sigma = y V}$$

Avoids forming XX^T

Avoids using inv(.)

Conditioning

unrw
regw] → "done right"

```
regw2 = np.linalg.solve(X.dot(X.transpose()) + ridge*np.eye(1+degree), Y.dot(X.transpose()))
unrw2 = np.linalg.solve(X.dot(X.transpose()), Y.dot(X.transpose()))
invw2 = np.linalg.inv(X.dot(X.transpose())).dot(Y.dot(X.transpose()))
print(f"difference in weights, ridge regression: {np.linalg.norm(regw-regw2):.4}")
print(f"difference in weights, plain regression: {np.linalg.norm(unrw-unrw2):.4}")
print(f"difference in weights, multiply by inv : {np.linalg.norm(unrw2-invw2):.4}")
```

→ 9×10^{-14}

→ 3.863

→ 2×10^{-4}

unrw – done right

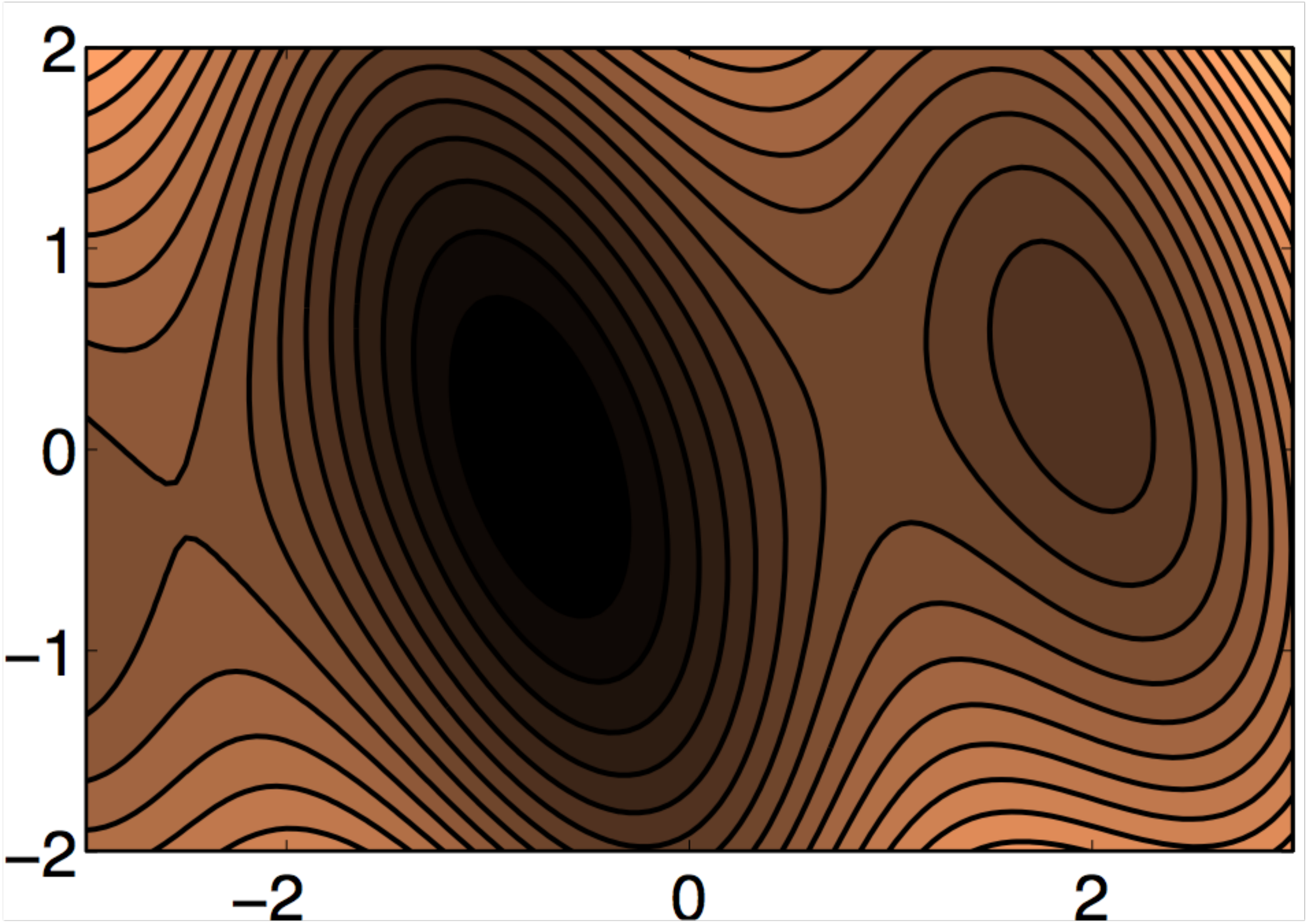
3.863 — ≠

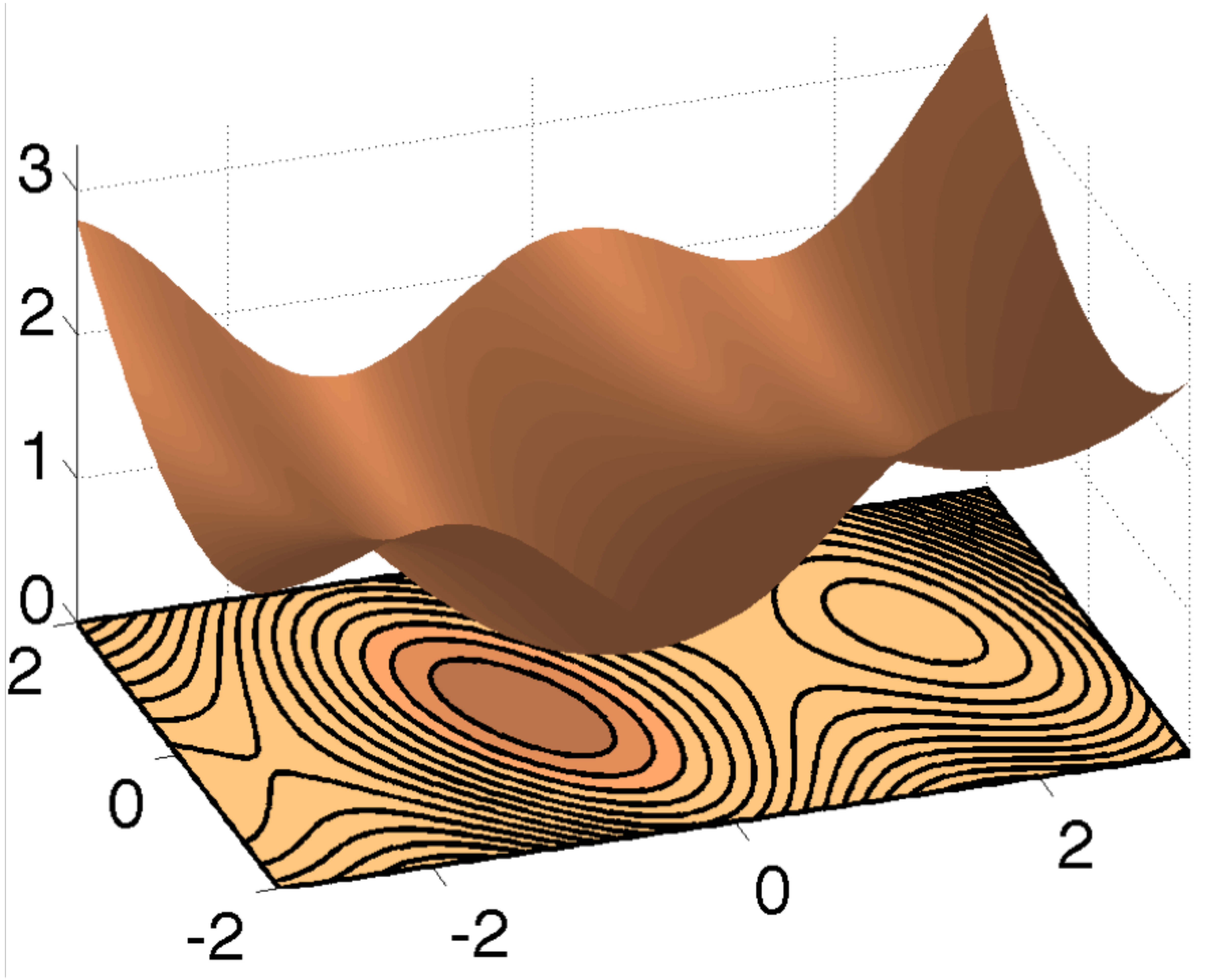
unrw2 – solve NE

2×10^{-4} — ≠

invw2 – use inv to solve NE

Contour plots





Calculus review

$$\frac{d}{dx}\left(3x^2 + x - 1\right) \rightarrow$$

$$3\frac{d}{dx}x^2 + \frac{d}{dx}x - \frac{d}{dx}1$$

$$\downarrow$$

$$2x$$

$$\frac{d}{dx}\sin x = \cos x$$

$$\frac{d}{dx}e^x = e^x$$

$$\frac{d}{dx}\left[(\sin x)\, e^x\right]$$

$$= \left[\frac{d}{dx}\sin x\right] e^x + (\sin x)\frac{d}{dx}e^x$$

$$f(g(x))$$

$$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx}$$

$$\sin x^2$$

$$(\cos x^2)(2x)$$

$$\frac{d}{dx}\left[4x^3 + (1+\sin x)^2 + (\cos x)(\sin x)\right]$$

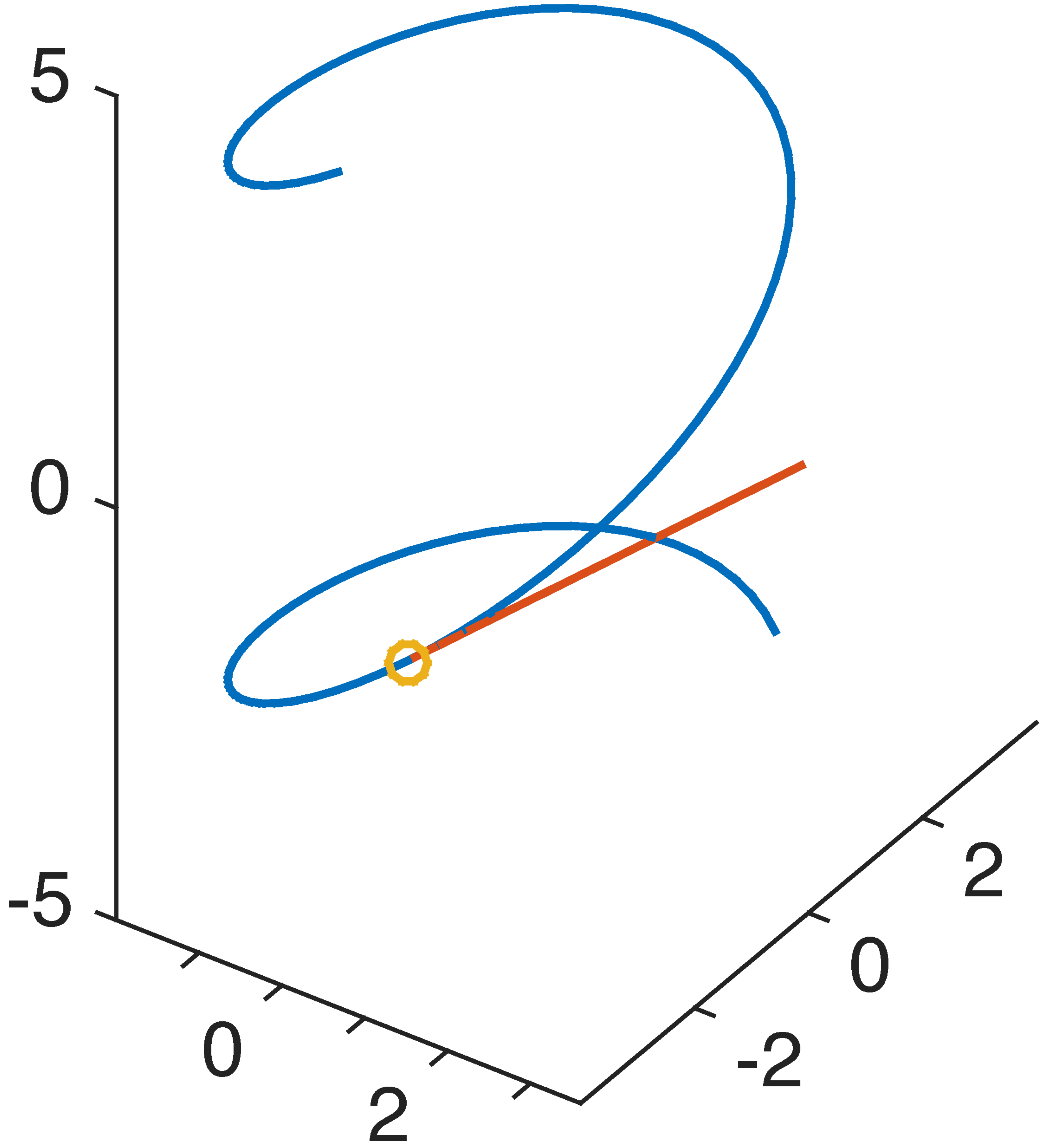
(enter in Canvas or email me if you're not on Canvas)

Derivative example

R -> R³

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}$$

$$\frac{d}{dx} f(x) = \begin{pmatrix} \frac{d}{dx} f_1(x) \\ \vdots \\ \frac{d}{dx} f_3(x) \end{pmatrix}$$



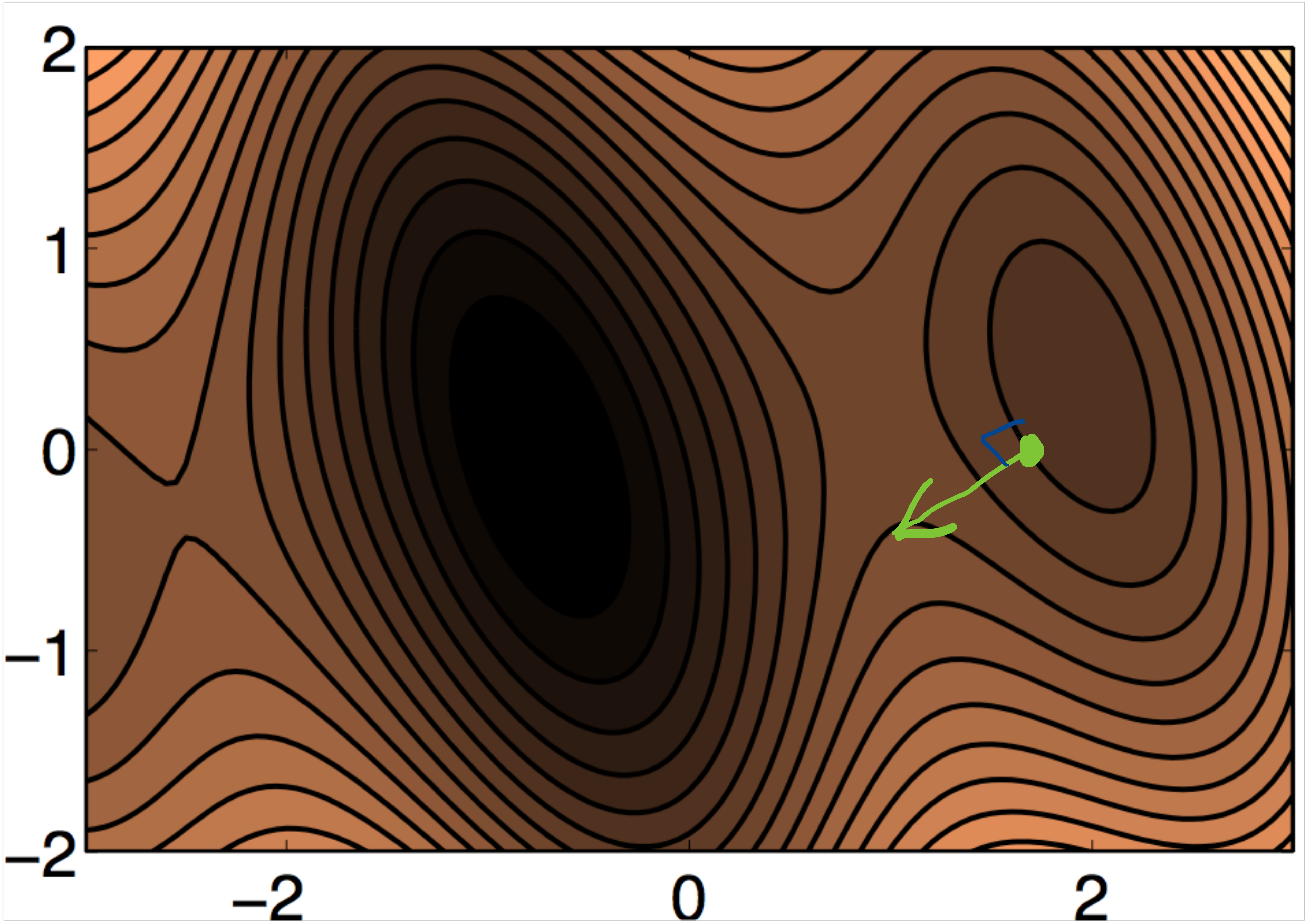
Derivative example

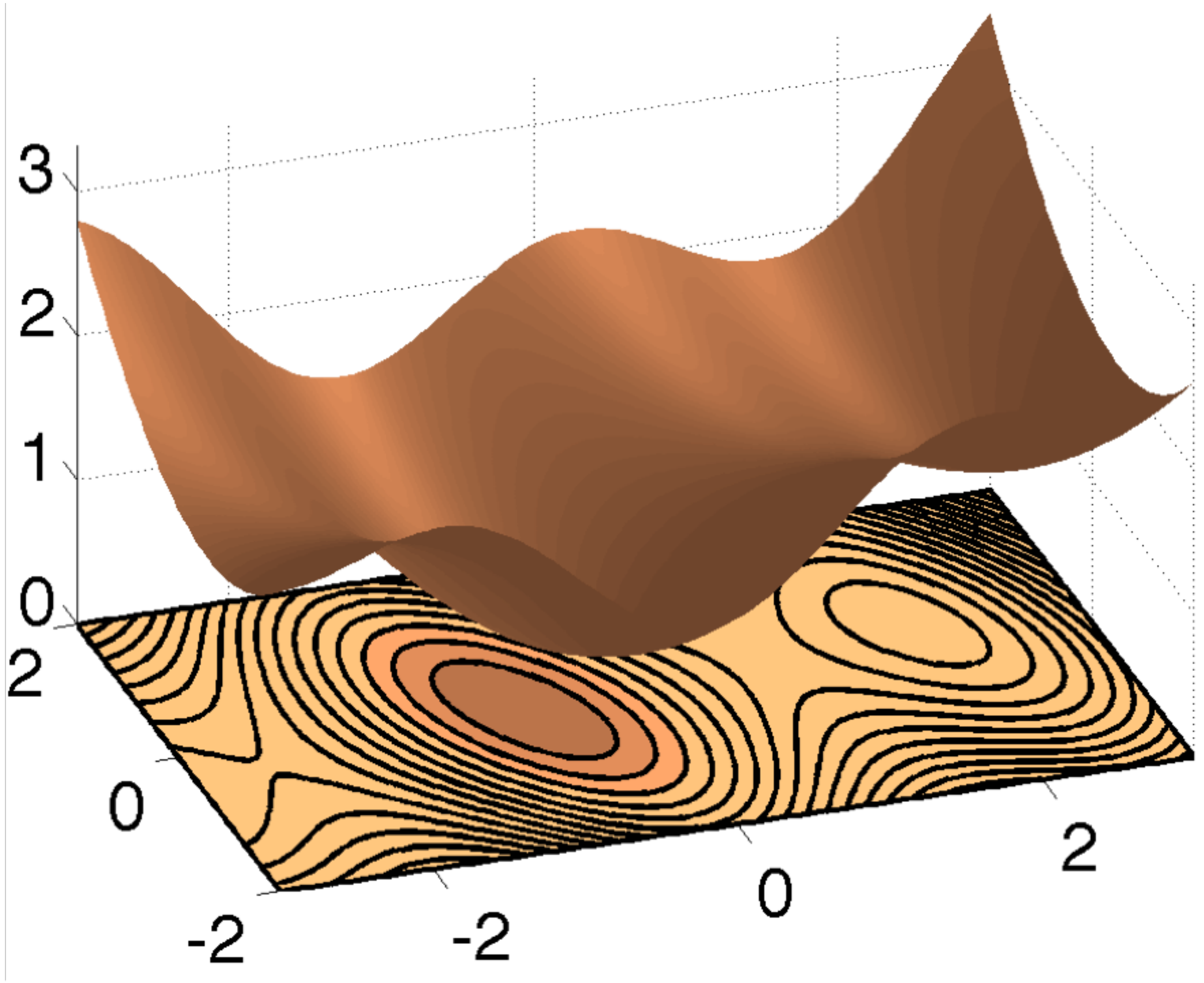
R^2 -> R

$f(x,y) \in \mathbb{R}$

$\left[\frac{\partial}{\partial x} f(x,y) \quad \frac{\partial}{\partial y} f(x,y)\right]$

$= \nabla f(x,y)^T$





coupling: $x = x(t)$
$y = y(t)$

$$u = \begin{pmatrix} x \\ y \end{pmatrix}$$

$f(x, y)$ where $x = x(t)$
$y = y(t)$

$$\frac{df}{dt} = \frac{df}{du}\frac{du}{dt}$$

$$= \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix}$$

$$= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$