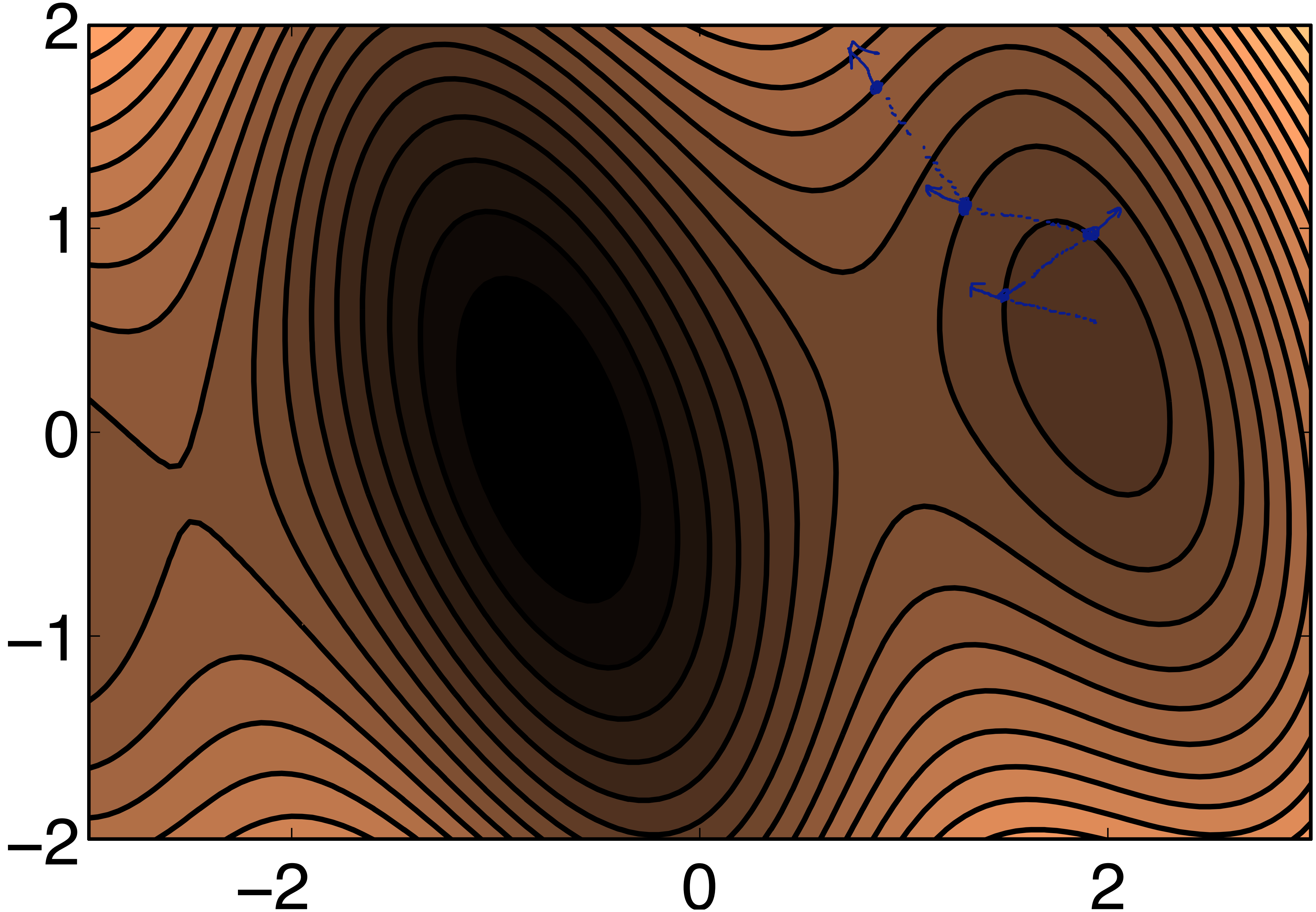
First-order methods
Convexity

10-725 Optimization
Geoff Gordon
Ryan Tibshirani

Gradient descent



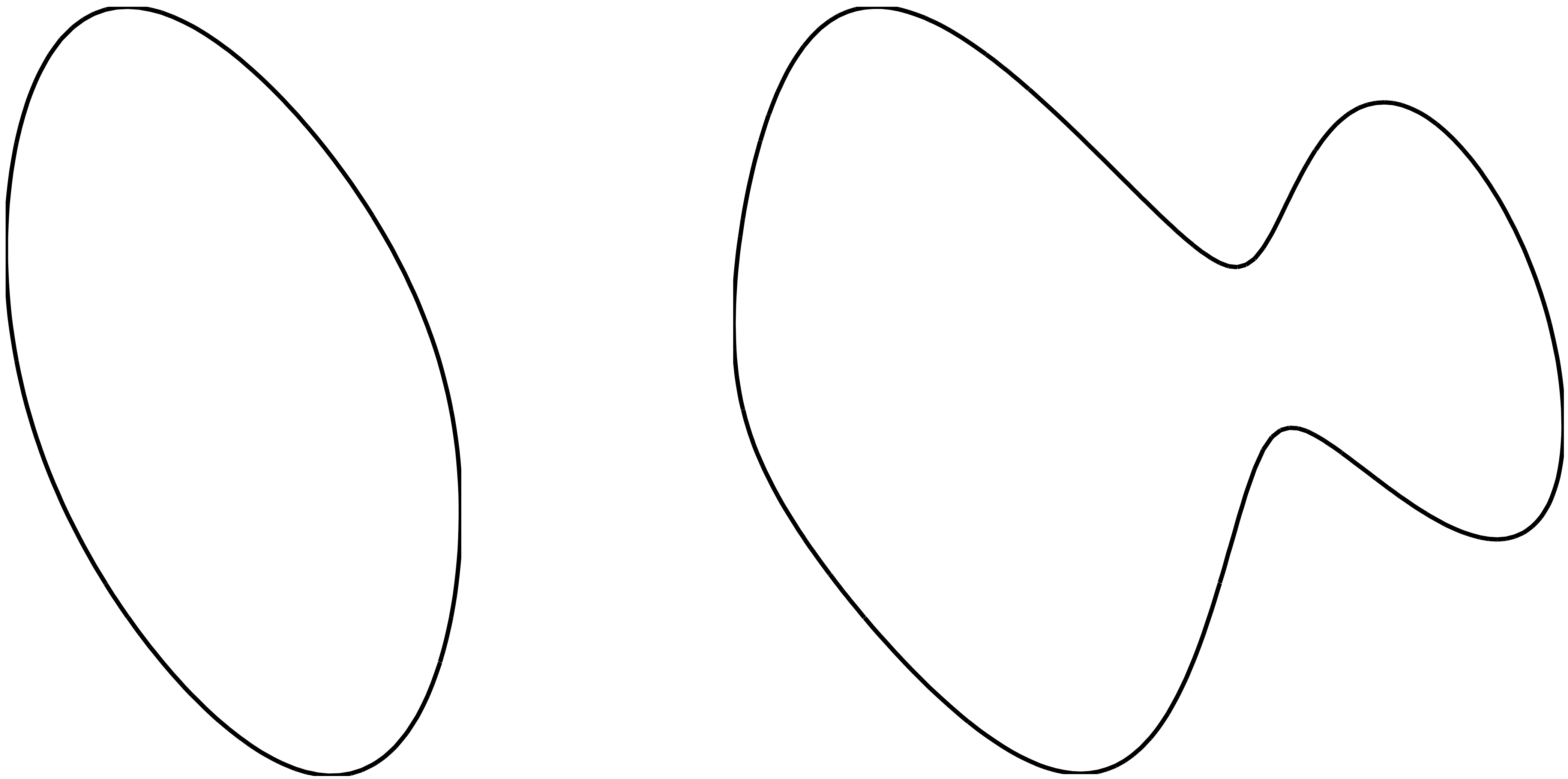
When do we stop?

- Using holdout set, if $f(x) = E(f_i(x) \mid i \sim P)$

$$E_{a,b}\,(a \cdot x - b)^2$$

- Using convergence bounds (later)
  - usual form is:
    - $K_f\,(f(x_0) - f(x^*))$ [some fn of $1/\epsilon$]
  - need estimates of first two terms
- For $f(x^*)$, duality (later); for K_f, properties of f:
  - convex? strongly convex? Lipschitz?

Convex sets



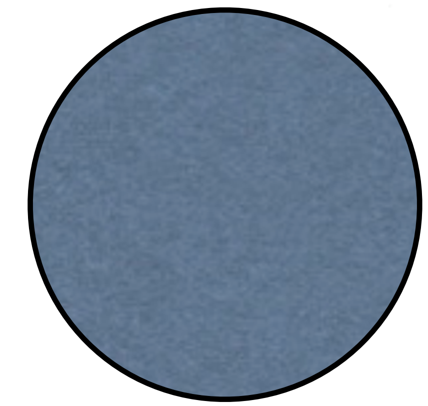
For all x, y ∈ C, for all t ∈ [0, 1]:

$$tx + (1-t)y \in C$$

Examples

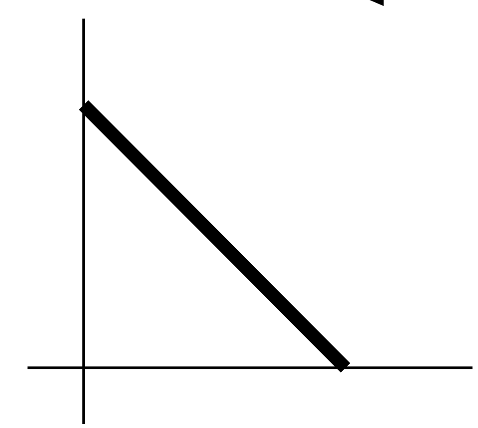
Boundaries

- x on boundary of C (∂C) if:
- x in interior of C if:
- x in *relative* interior (rel int C) if:
- C closed if:
- C open if:
- C compact if:



$\{x \mid ||x|| \leq 1\}$

<

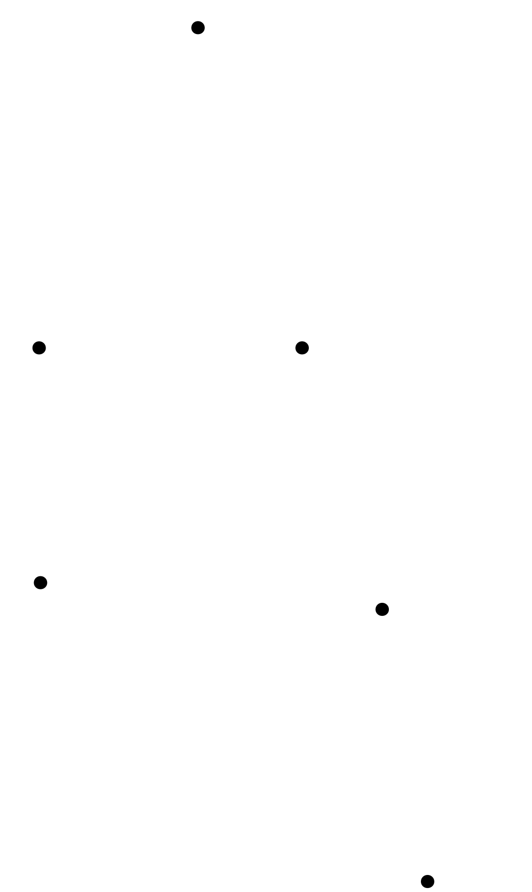


$\{x \mid x \geq 0, x_1 + x_2 = 1\}$

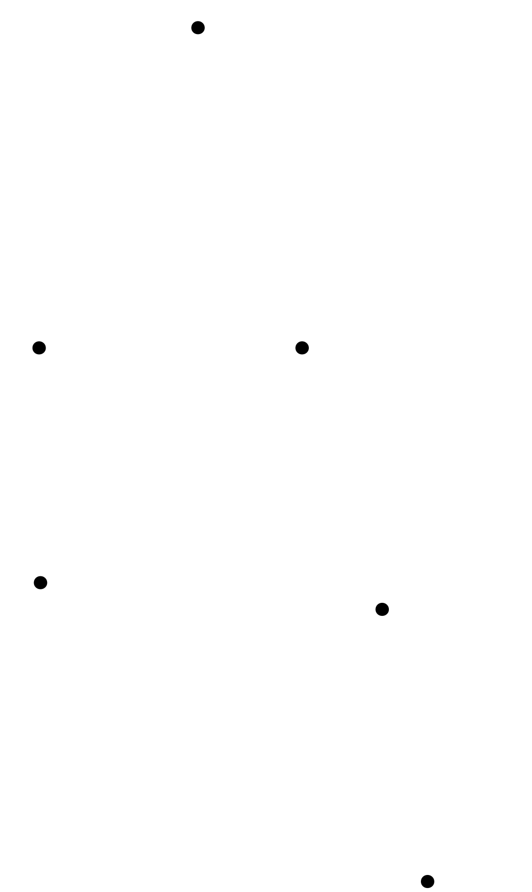
R^n

$\varnothing$

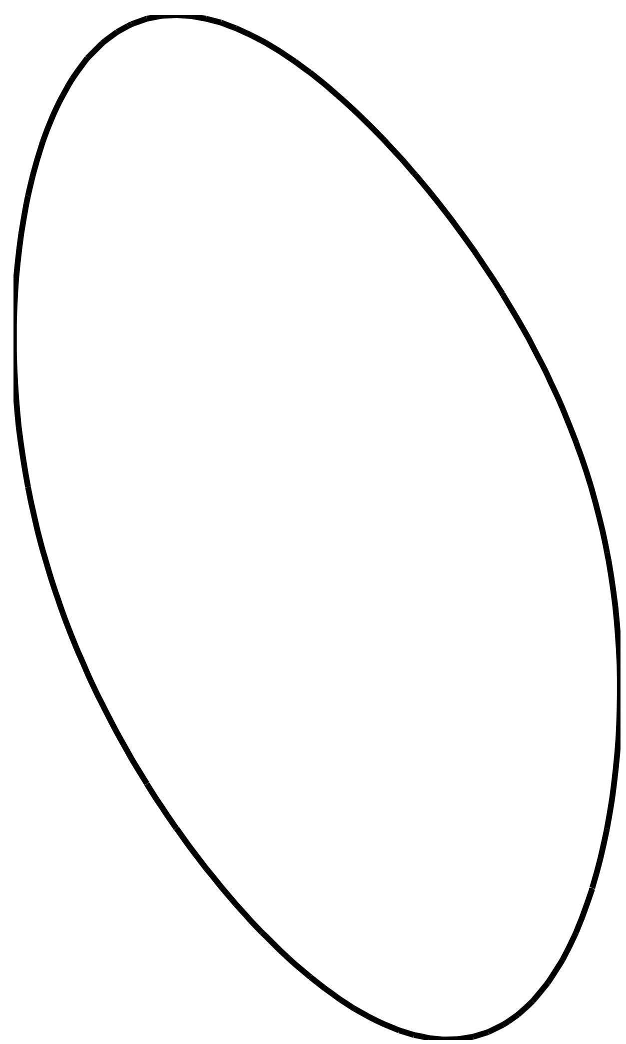
Convex hull



Dual representation



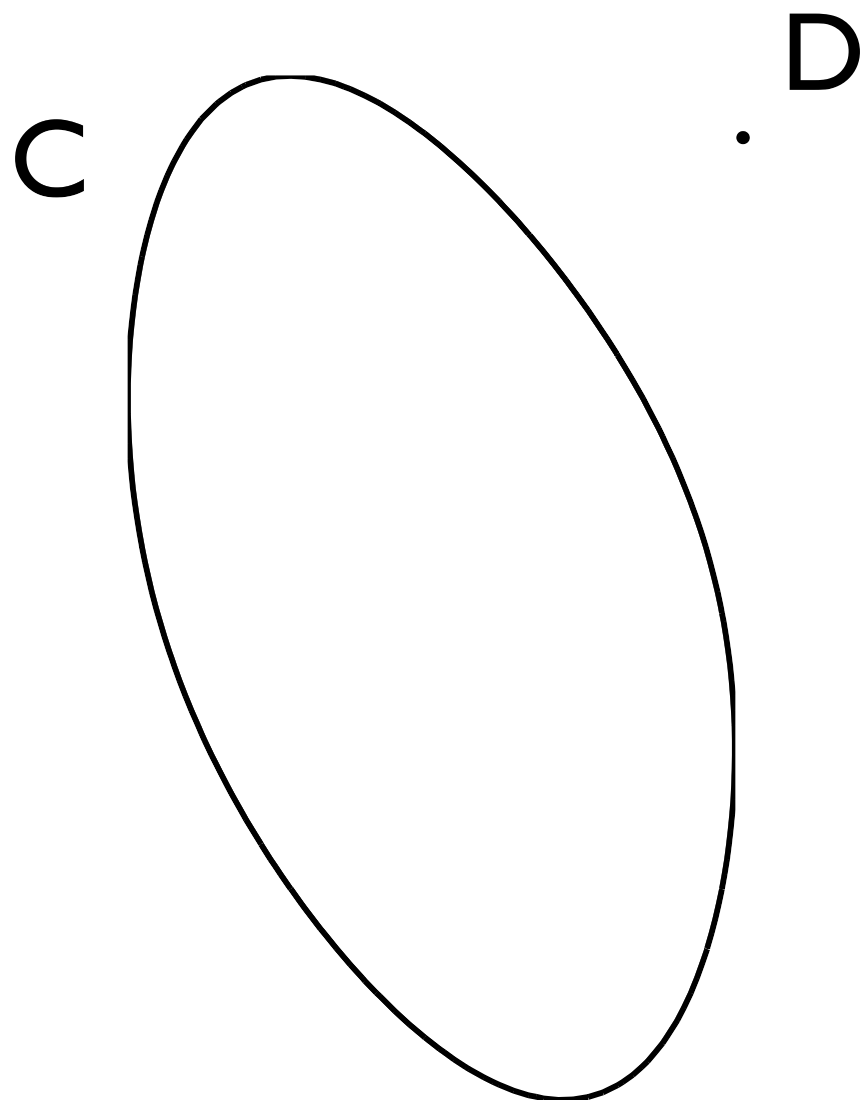
Supporting hyperplane thm



- For any point x_0 on boundary of convex C:
  - exist (w, b) with
  -
  -

Supporting hyperplane exs

Separating hyperplane thm



- For any convex C and D with
  - exist (w, b) with
  -
  -
- If both C, D are closed, and at least one compact:

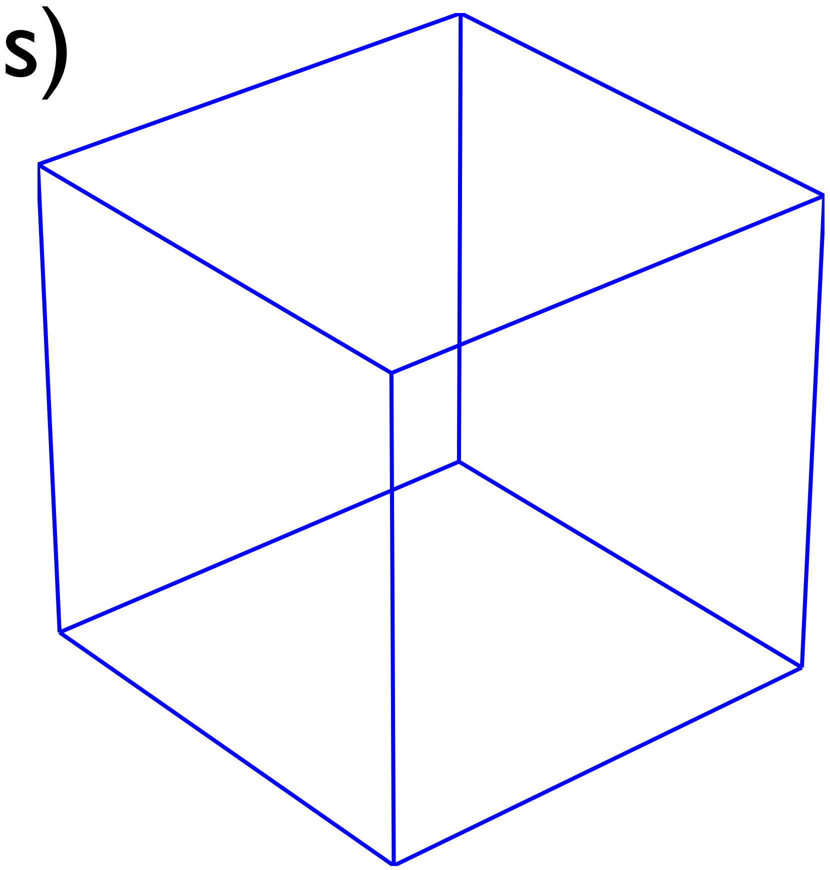
Separating hyperplane exs

Proving a set convex

- Use definition directly
- Represent as convex hull or ∩ of halfspaces
- Supporting hyperplane partial converse
  - C closed, nonempty interior, has supporting hyperplane at all boundary points ⇒ C convex
- Build C up from simpler sets using convexity-preserving operations

Convexity-preserving set ops

- Translation
- Scaling
- Affine fn
  - projection (e.g., dropping coords)
- Intersection
- Set sum
- Perspective



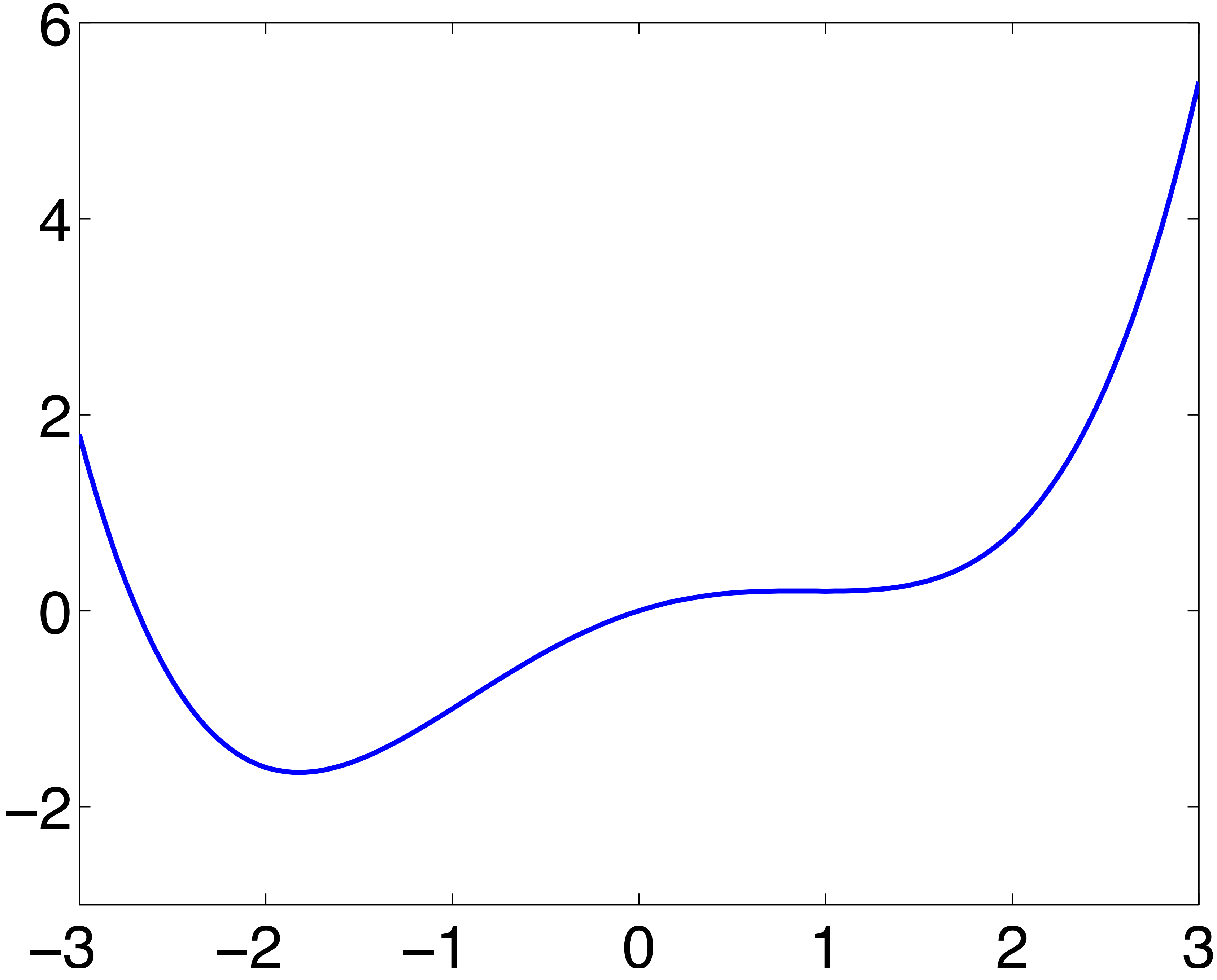
Ex: symmetric PSD matrices

- Two proofs that $\{A \mid A = A^T, A \succcurlyeq 0\}$ is convex
  - $x^T (tA + (1-t)B) x =$
  - $x^TAx =$

Ex: conditionals

- Given a convex set of dist'ns $P(x_{1:7})$,
  - $P(x_{1:5} \mid x_{6:7}) =$
  - numerator:
  - denominator:
- Convex?

Epigraph

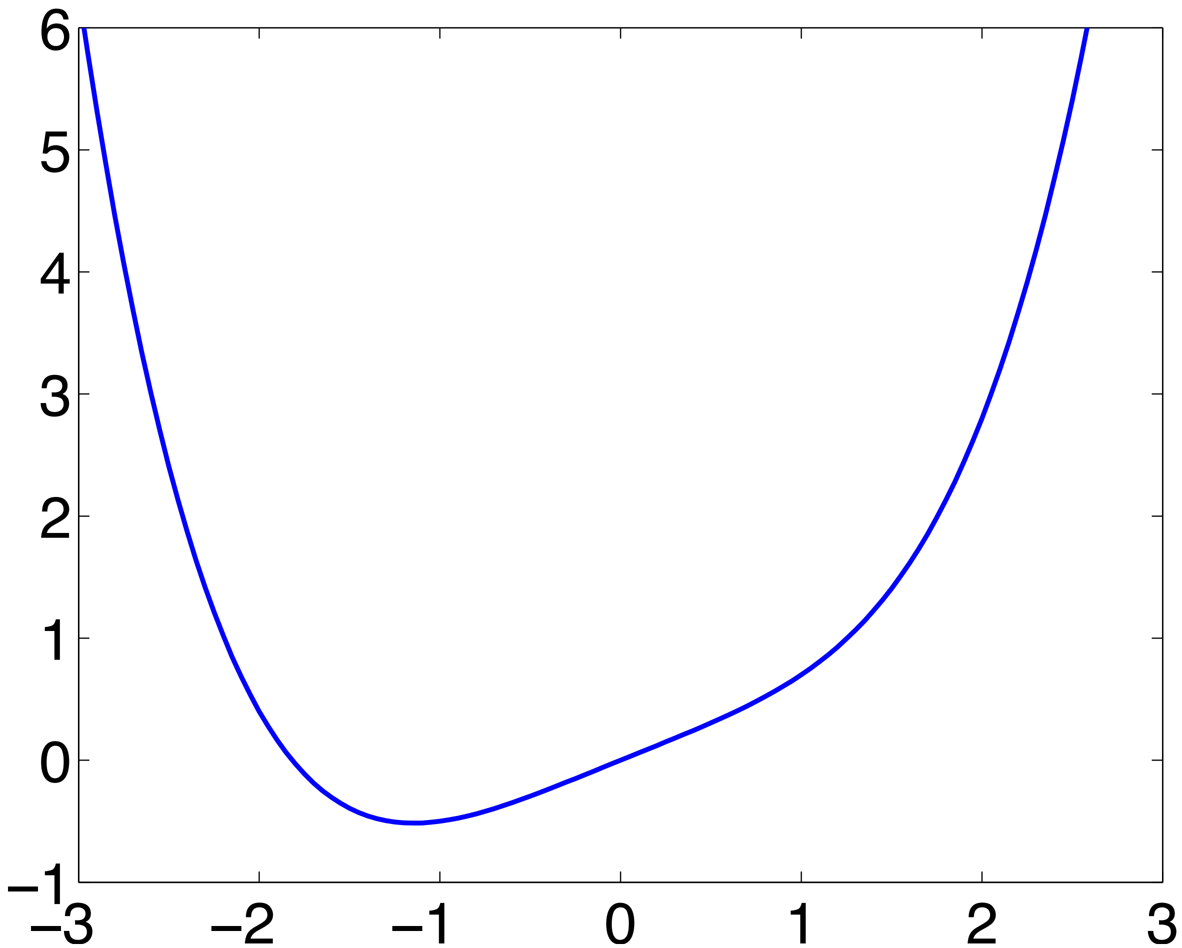


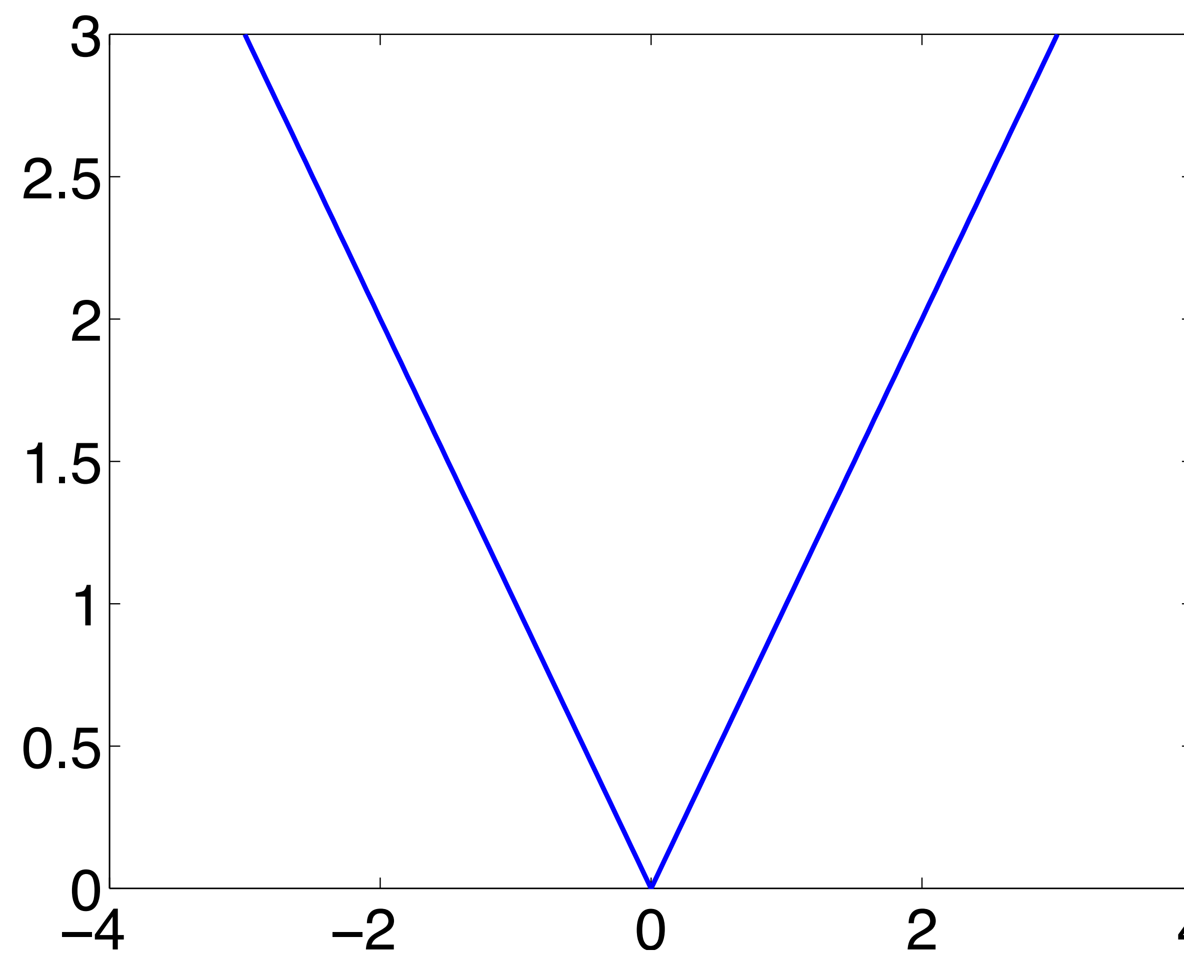
Domain

- dom f =

- dom 1/x =
- dom ln(x) =

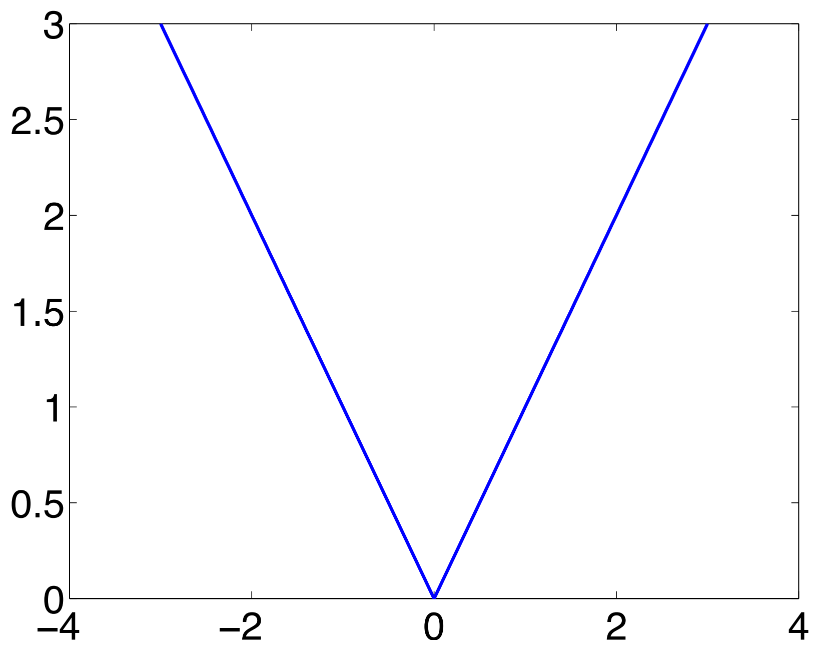
Convex functions





Relating convex sets and fns

- f(x) convex ⇒ { x | } convex
- Converse?



Proving a function convex

- Use definition directly
- Prove that epigraph is convex via set methods
  - e.g., supporting hyperplanes: for all x, y,
  - this is first-order convexity condition for fns
- 2nd order:
- Construct f from simpler convex fns using convexity-preserving ops

Convexity-preserving fn ops

- Nonnegative weighted sum
- Pointwise max/sup
- Composition w/ affine
- Composition w/ monotone convex
- Perspective
- f(x, y) convex in (x, y), set C convex:
  - $g(y) = \min_{x \in C} f(x,y)$ is convex if $g(y) > -\infty$

Example: $f(x) = |x|$

In 2 or more dimensions

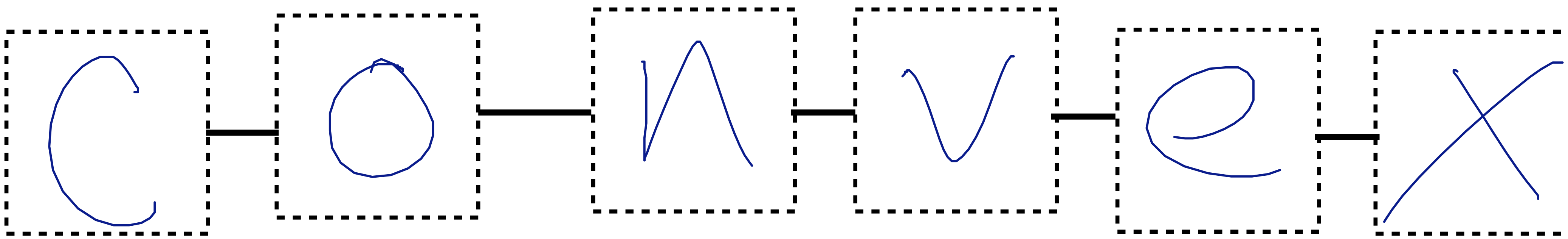
- All the above, but for 2nd order:

- Or: reduce to 1D

Ex: structured classifier

c o n v e x

Ex: structured classifier



x_i

y_i

$\varphi_j(x_i)$

$\psi_{ijk}(x_i, y_i)$

$\chi_{ikl}(y_i, y_{i+1})$

$L(x,y;v,w) =$

Classifier:

Learning structured classifier

- Get it right if:
- So, want:
- Where $\pi(y,y') = \{$
- RHS:
- RHS – LHS:
- Train: lots of pairs (x^t,y^t)