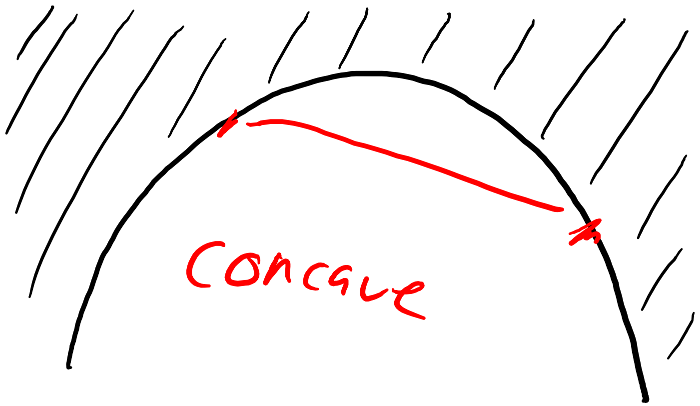
CPSC 340:
Machine Learning and Data Mining

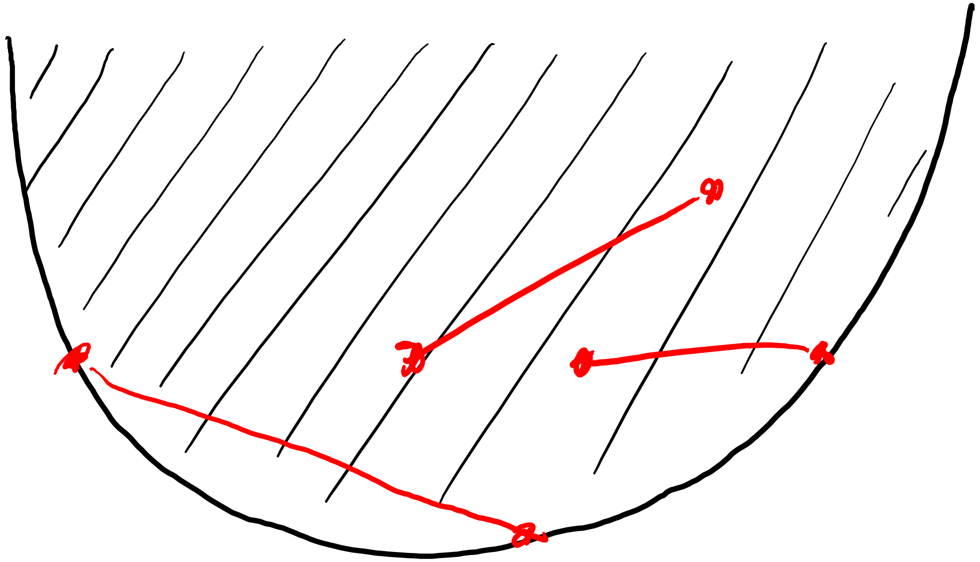
Notes on Convex Functions

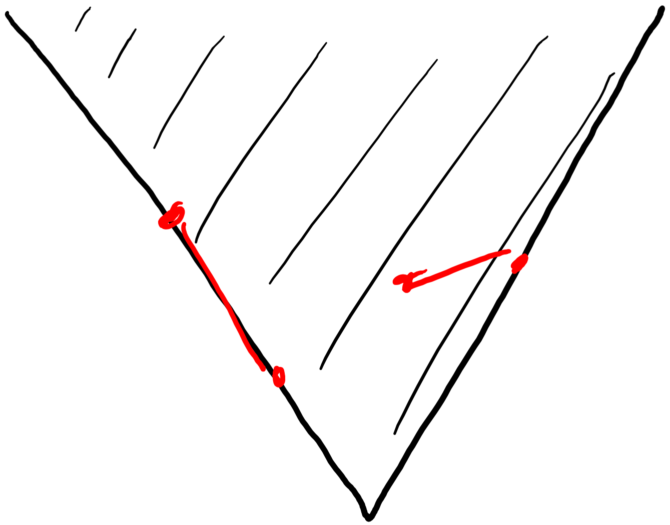
Original version of these slides by Mark Schmidt, with modifications by Mike Gelbart.

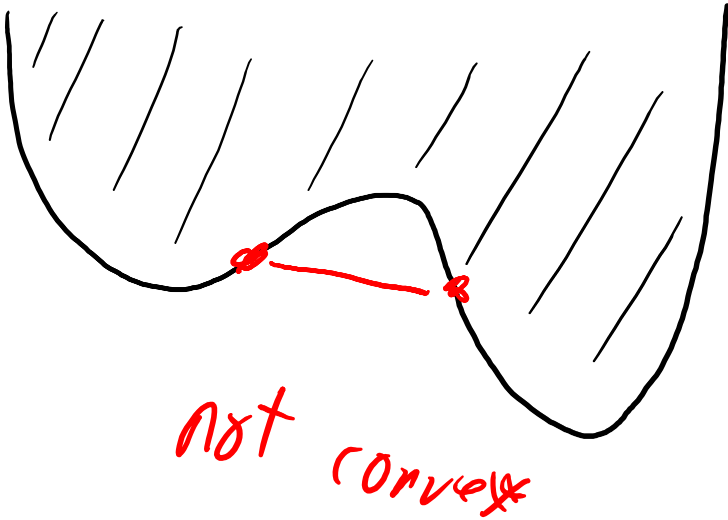
Convex Functions



- Is finding a ‘w’ with $\nabla f(w) = 0$ good enough?
  - Yes, for convex functions.







- A function is convex if the area above the function is a convex set.
  - All values between any two points above function stay above function.

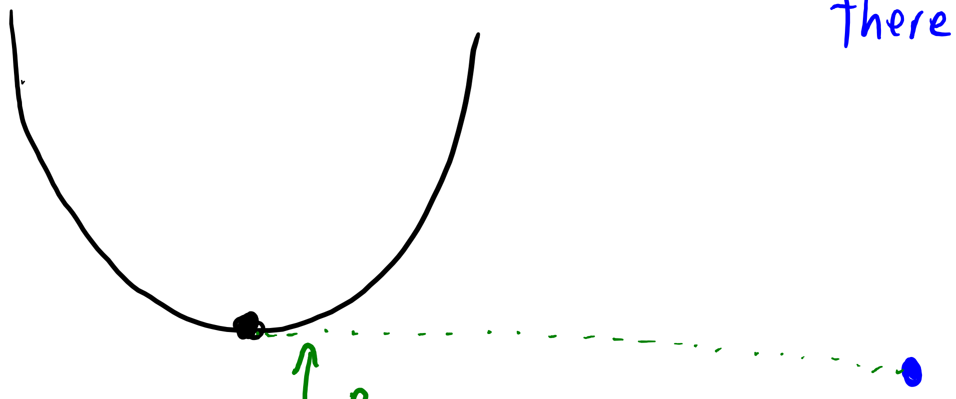
Convex Functions

- All 'w' with ∇ f(w) = 0 for convex functions are global minima.

Proof by contradiction:

Consider a local minimum

If this is not global minimum, there must a smaller value.



By convexity we can move along line to global minimum and decrease objective.

But this contradicts that we are at a local minimum.

  - Normal equations finds a global minimum because of convexity.

How do we know if a function is convex?

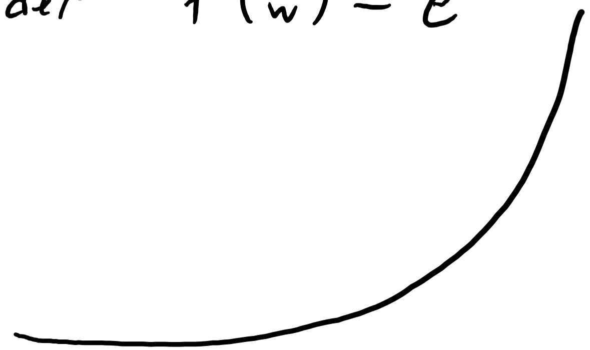
- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all 'w'.

Consider $f(w) = \frac{1}{2}aw^2$ for $a > 0$.

We have $f'(w) = aw$
and $f''(w) = a > 0$

By assumption

Consider $f(w) = e^w$



We have $f'(w) = e^w$
and $f''(w) = e^w > 0$

By definition of exponential function.

How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all 'w'.
  - A convex function multiplied by non-negative constant is convex.

We showed that $f(w) = e^w$ is convex, so $f(w) = 10e^w$ is convex.

How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff f''(w) ≥ 0 for all 'w'.
  - A convex function multiplied by non-negative constant is convex.
  - Norms and squared norms are convex.

$\|w\|$, $\|w\|^2$, $\|w\|_1$, $\|w\|_\infty$, $\|w\|_1^2$, and so on are all convex.

How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all 'w'.
  - A convex function multiplied by non-negative constant is convex.
  - Norms and squared norms are convex.
  - The sum of convex functions is a convex function.

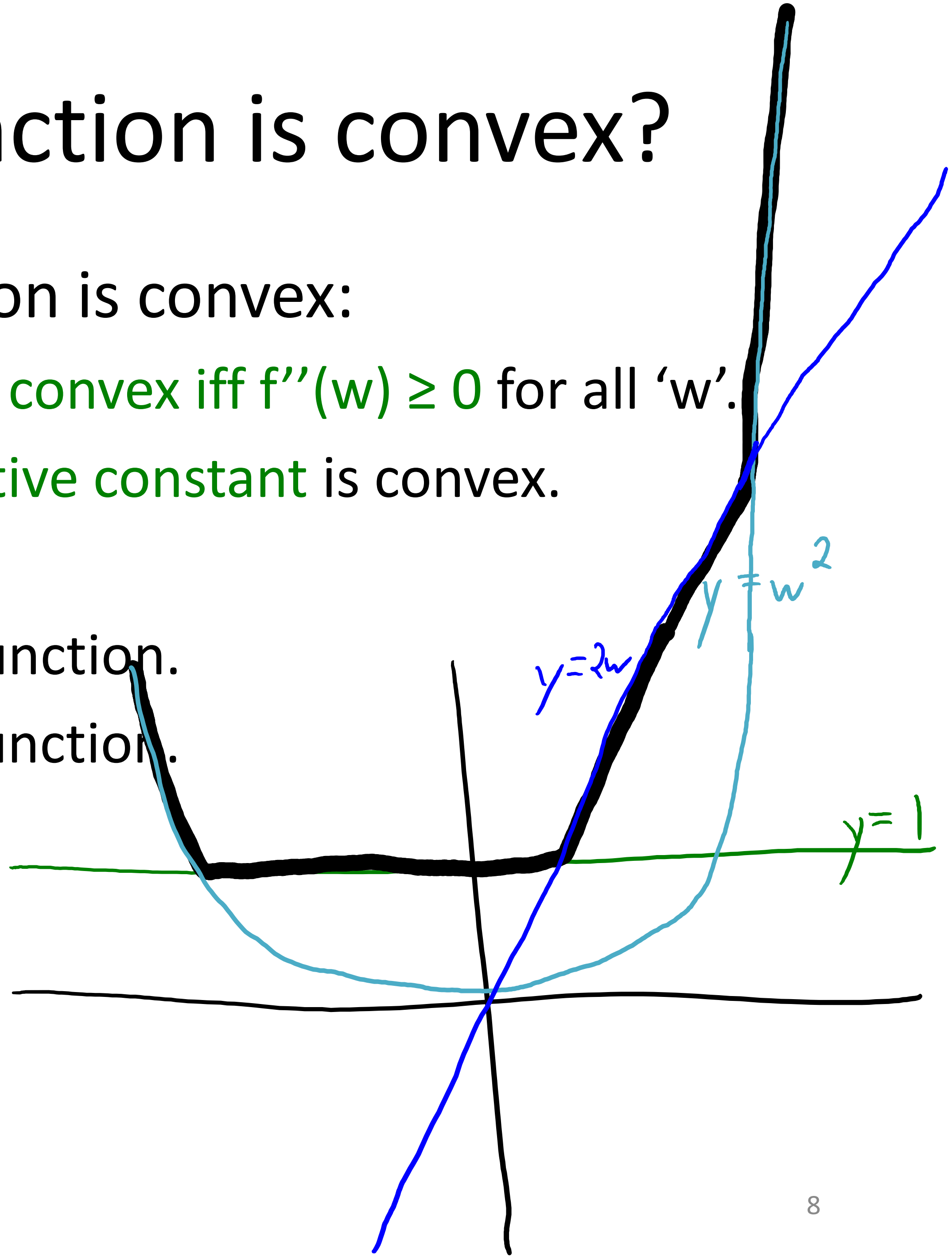
$$f(x) = \underbrace{10e^{w}}_{\text{From earlier}} + \underbrace{\frac{\lambda}{2}}_{\text{constant}}\underbrace{\|w\|^2}_{\text{norm squared}} \quad \text{is convex}$$

How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all ‘w’.
  - A convex function multiplied by non-negative constant is convex.
  - Norms and squared norms are convex.
  - The sum of convex functions is a convex function.
  - The max of convex functions is a convex function.

$f(w) = \max\{1, 2w, w^2\}$ is convex.

convex



How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all ‘w’.
  - A convex function multiplied by non-negative constant is convex.
  - Norms and squared norms are convex.
  - The sum of convex functions is a convex function.
  - The max of convex functions is a convex function.
  - Composition of a convex function and a linear function is convex.

If ‘f’ is convex the $f(Xw - y)$ is convex.

($Xw - y$: linear function)

How do we know if a function is convex?

- Some useful tricks for showing a function is convex:
  - 1-variable, twice-differentiable function is convex iff $f''(w) \geq 0$ for all 'w'.
  - A convex function multiplied by non-negative constant is convex.
  - Norms and squared norms are convex.
  - The sum of convex functions is a convex function.
  - The max of convex functions is a convex function.
  - Composition of a convex function and a linear function is convex.
- But: not true that composition of convex with convex is convex:

Even if 'f' is convex and 'g' is convex, $f(g(w))$ might <u>not</u> be convex.

E.g. x^2 is convex and $-\log(x)$ is convex but $-\log(x^2)$ is <u>not</u> convex.