Convex Optimization

CS5491: Artificial Intelligence
ZHICHAO LU

Content Credits: **Prof. Wei**'s CS4486 Course and **Prof. Boddeti**'s AI Course

TODAY

Convexity

Convex Optimization

CONVEX OPTIMIZATION: DEFINITION

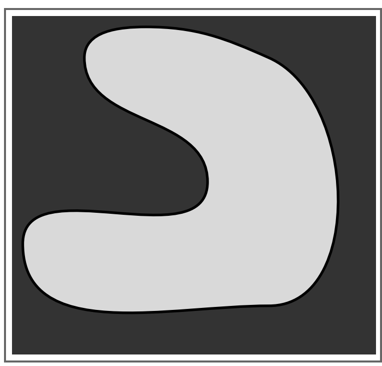
Convex Optimization Problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F} \end{aligned}$$

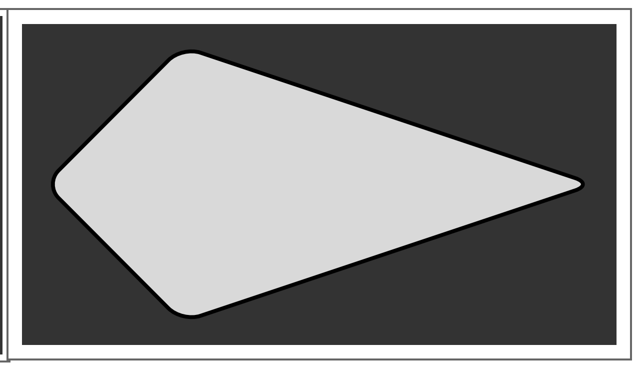
A special class of optimization problem

An optimization problem whose optimization objective f is a convex function and feasible region $\mathcal{F}$ is a convex set.

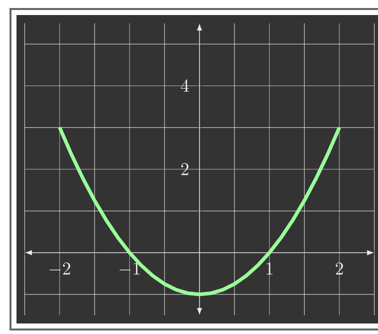
Nonconvex Set



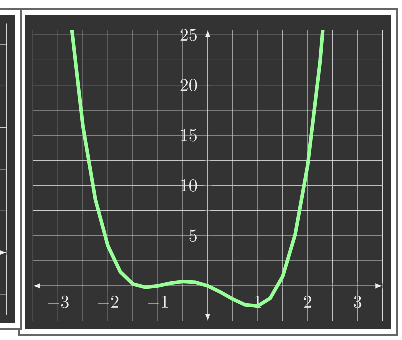
Convex Set



Convex Function



NonConvex Function



CONVEX COMBINATION

A point between two points

Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, a convex combination of them is any point of the form $\mathbf{z} = \theta\mathbf{x} + (1-\theta)\mathbf{y}$ where $\theta \in [0, 1]$.

When $\theta \in (0, 1)$, $\mathbf{z}$ is called a strict convex combination of $\mathbf{x}$, $\mathbf{y}$.

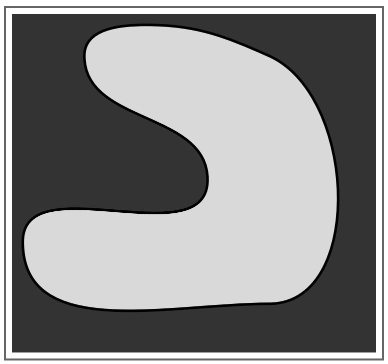
CONVEX SETS

Conceptually: Any convex combination of two points in the set is also in the set

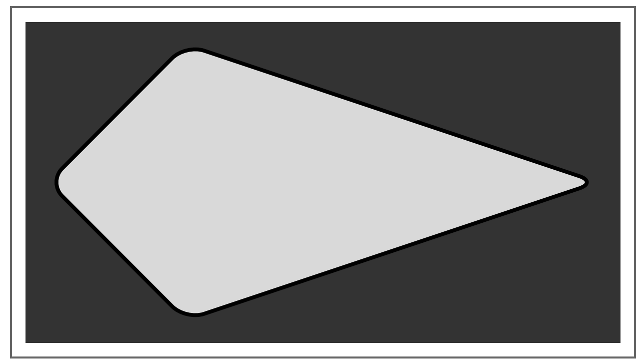
Mathematically: A set $\mathcal{F}$ is convex if $\forall x, y \in \mathcal{F}, \forall \theta \in [0, 1]$,

$$z = \theta x + (1 - \theta) y \in \mathcal{F}$$

Nonconvex Set



Convex Set



QUIZ: CONVEX SET

Which of the following sets are convex?

- $\mathcal{F} = \mathbb{R}^n$
- $\mathcal{F} = \emptyset$
- $\mathcal{F} = \{\mathbf{x}_0\}, \mathbf{x}_0 \in \mathbb{R}^n$
- $\mathcal{F} = \mathcal{F}_1 \bigcap \mathcal{F}_2$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are convex sets.
- $\mathcal{F} = \mathcal{F}_1 \bigcup \mathcal{F}_2$, where $\mathcal{F}_1$ and $\mathcal{F}_2$ are convex sets.
- $\mathcal{F} = \mathbb{Z}^n$

CONVEX FUNCTION

Value in the middle point is lower than average value

Let $\mathcal{F}$ be a convex set. A function $f : \mathcal{F} \to \mathbb{R}$ is convex in $\mathcal{F}$ if $\forall x, y \in \mathcal{F}, \forall \theta \in [0, 1]$,

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y})$$

If $\mathcal{F} = \mathbb{R}^n$, we simply say f is convex.

HOW TO DETERMINE IF A FUNCTIONS IS CONVEX?

Prove by definition

Use properties of convex functions

- Sum of convex functions is convex

$$\text{If } f(\mathbf{x}) = \sum_i w_i f_i(\mathbf{x}), w_i \geq 0, f_i(\mathbf{x}) \text{ convex, then } f(\mathbf{x}) \text{ is convex.}$$

- Convexity is preserved under a linear transformation

$$\text{If } f(\mathbf{x}) = g(A\mathbf{x} + b), \text{ if } g(\mathbf{x}) \text{ is convex, then } f(\mathbf{x}) \text{ is convex.}$$

- If f is a twice differentiable function of one variable, f is convex on an interval $[a, b] \subseteq \mathbb{R}$ iff (if and only if) its second derivative $f''(x) \geq 0$ in $[a, b]$

CONVEX OPTIMIZATION: DEFINITION

If f is a twice continuously differentiable function of n variables, f is convex on $\mathcal{F}$ iff its Hessian matrix of second partial derivatives is positive semidefinite on the interior of $\mathcal{F}$.

$$H(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 x_n} & \frac{\partial^2 f}{\partial x_2 x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

H is positive semidefinite in S if $\forall \mathbf{x} \in S, \forall \mathbf{z} \in \mathbb{R}^n, \mathbf{z}^T H(\mathbf{x})\mathbf{z} \geq 0$

H is positive semidefinite in $\mathbb{R}^n$ iff all eigenvalues of H are non-negative.

Alternatively, prove $\mathbf{z}^T H(\mathbf{x})\mathbf{z} = \sum_i g_i^2(x, z)$

CONCAVITY AND CONVEXITY

Concave function

- A function f is concave if $-f$ is convex
- Let $\mathcal{F}$ be a convex set. A function $f : \mathcal{F} \rightarrow \mathbb{R}$ is concave in $\mathcal{F}$ if $\forall \mathbf{x}, \mathbf{y} \in \mathcal{F}, \forall \theta \in [0, 1]$,

$$f(\theta \mathbf{x} + (1 - \theta)\mathbf{y}) \geq \theta f(\mathbf{x}) + (1 - \theta) f(\mathbf{y})$$

The following is a convex optimization problem if f is a concave function and $\mathcal{F}$ is a convex set

$$\begin{aligned} \max_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{F} \end{aligned}$$

QUIZ 2: CONVEX FUNCTION

Which of the following functions are convex?

- $f(x) = e^{ax}, a \in \mathbb{R}$
- $f(x) = \log x, x \in (0, +\infty)$
- $f(x) = \|\mathbf{x}\|_2 = \sqrt{\sum_i^n x_i^2}$
- $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y}, \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$
- $f(x) = x^3, x \in \mathbb{R}$

CONVEX OPTIMIZATION: LOCAL OPTIMA=GLOBAL OPTIMA

$$\min_{\mathbf{x}} \quad f(\mathbf{x})$$
$$\text{s.t.} \quad \mathbf{x} \in \mathcal{F}$$

Given an optimization problem, a point $\mathbf{x} \in \mathbb{R}^n$ is globally optimal if $\mathbf{x} \in \mathcal{F}$ and $\forall \mathbf{y} \in \mathcal{F}$, $f(\mathbf{x}) \leq f(\mathbf{y})$

Given an optimization problem, a point $\mathbf{x} \in \mathbb{R}^n$ is locally optimal if $\mathbf{x} \in \mathcal{F}$ and $\exists R > 0$ such that $\forall \mathbf{y} : \mathbf{y} \in \mathcal{F}$ and $\|\mathbf{x} - \mathbf{y}\|_2 \leq R$, $f(\mathbf{x}) \leq f(\mathbf{y})$

Theorem 1: For a convex optimization problem, all locally optimal points are globally optimal

OPTIMIZATION

CONVEX OPTIMIZATION: HOW TO SOLVE?

Recall discrete setting

- Local search
- Iteratively improving an assignment

Continuous and differentiable setting

- Iteratively improving value of $\mathbf{x}$
- Based on gradient

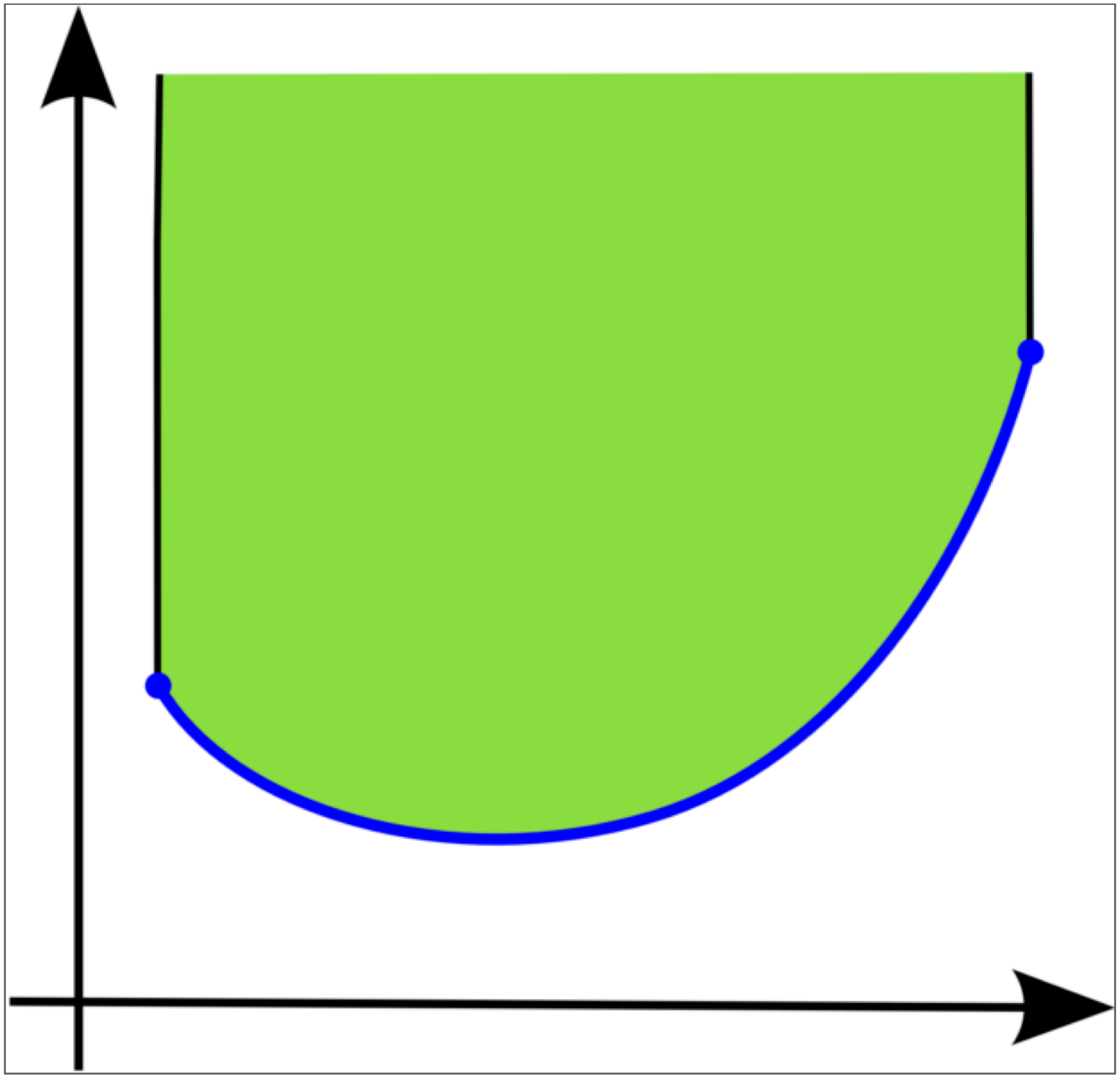
CONVEX OPTIMIZATION: HOW TO SOLVE?

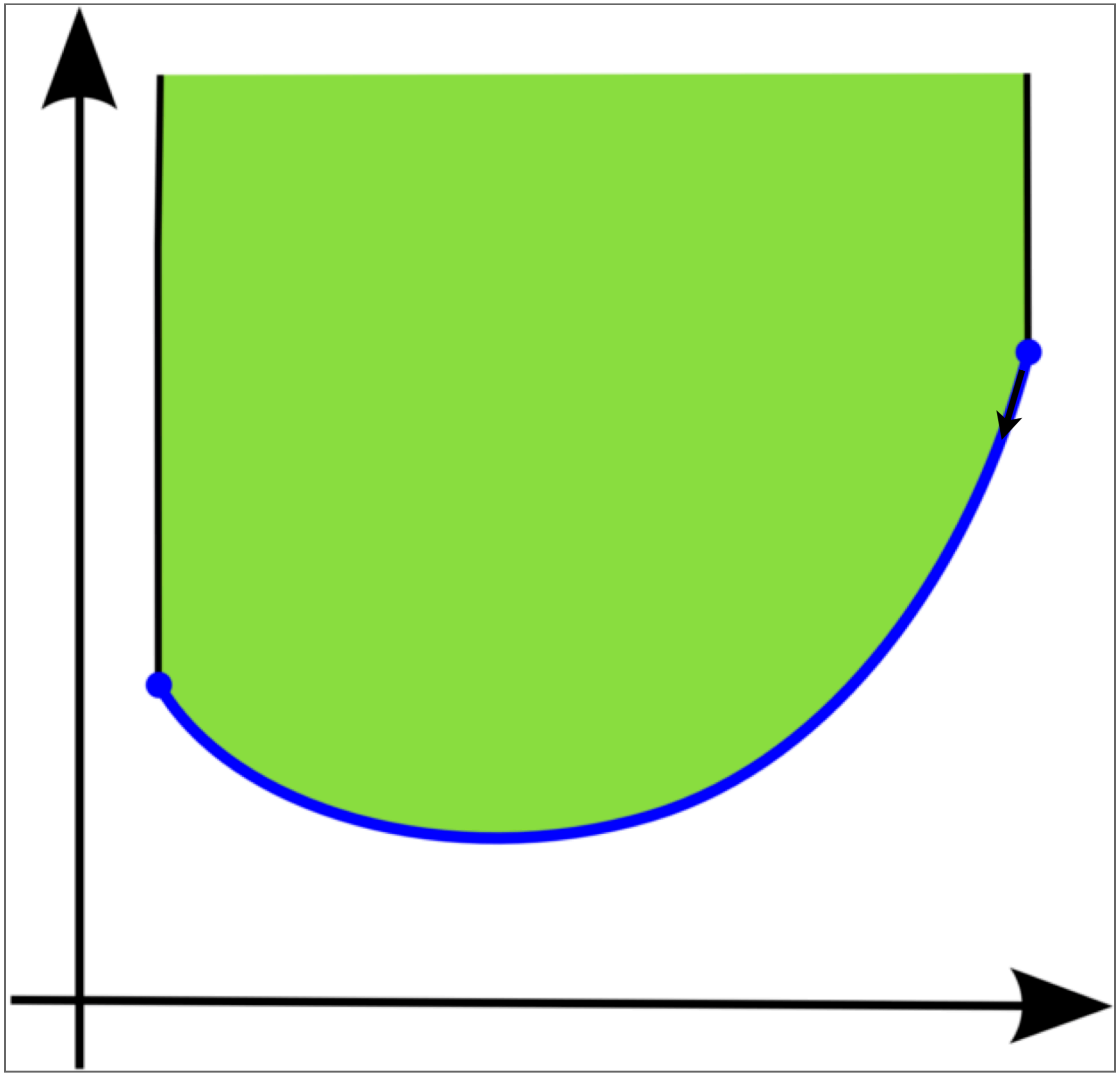
For $f : \mathbb{R}^n \to \mathbb{R}$, gradient is the vector of partial derivatives

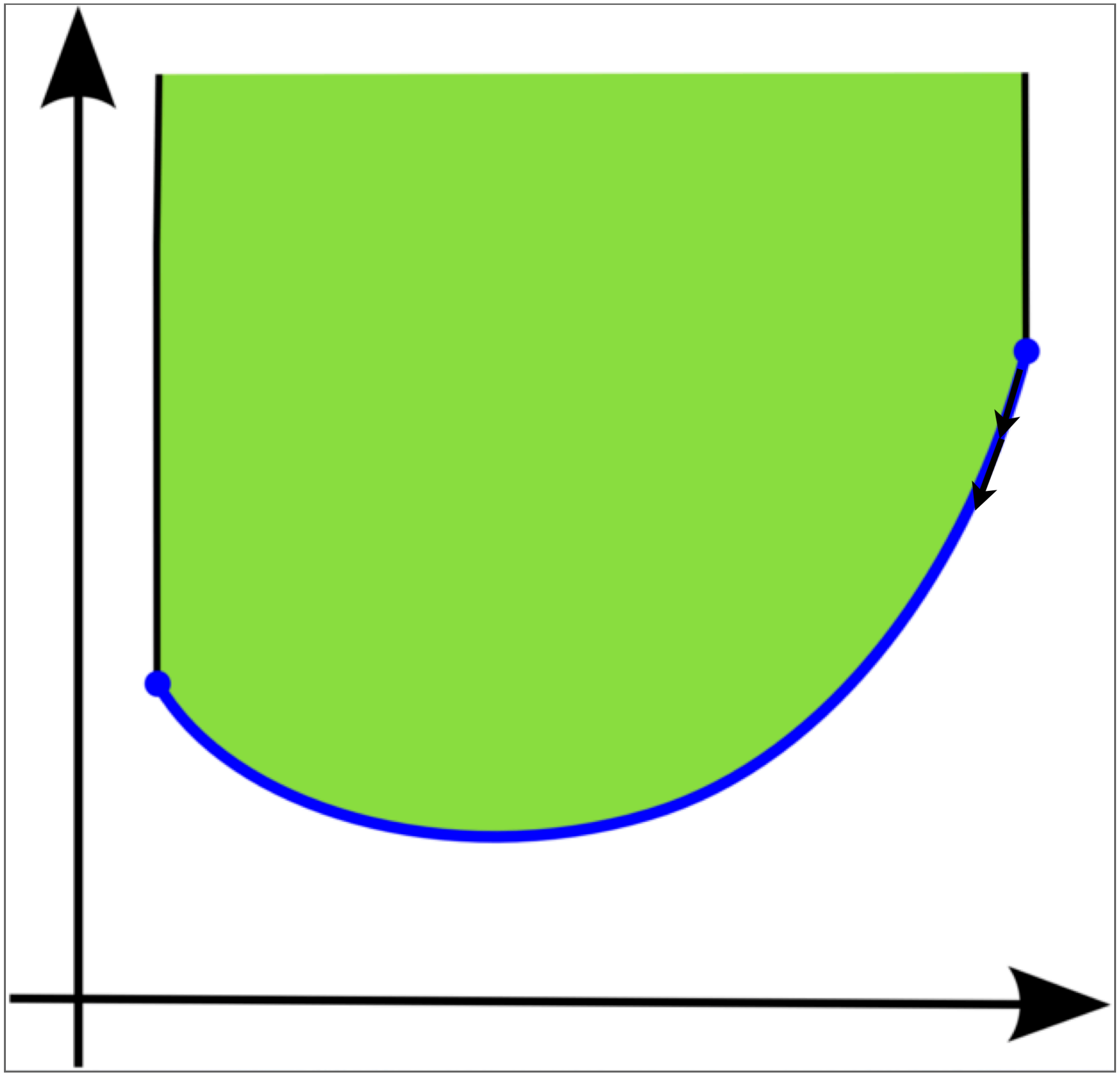
- A multi-variable generalization of the derivative
- Point in the direction of steepest increase in f

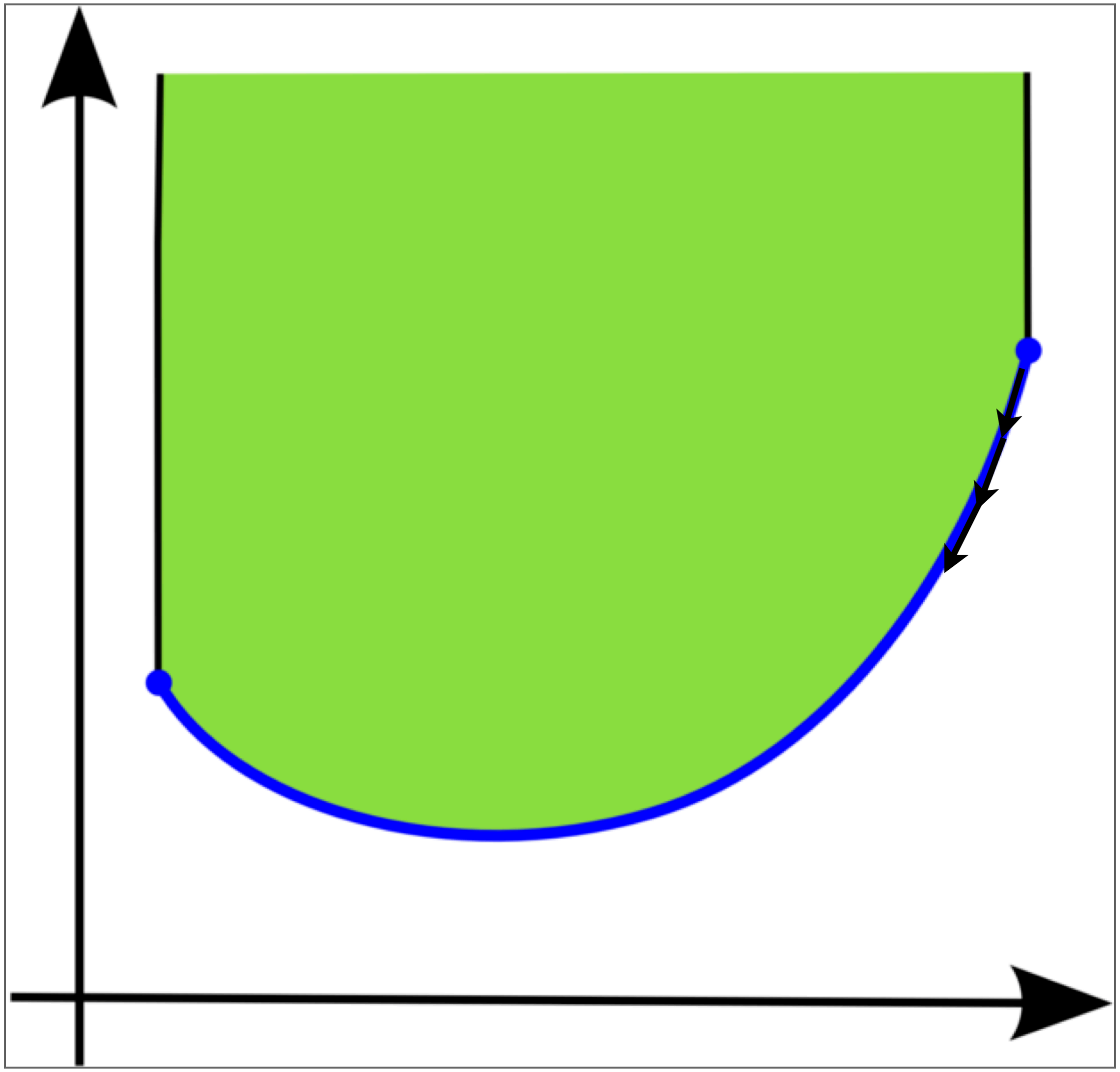
$$\bigtriangledown f(\mathbf{x}) \in \mathbb{R}^n = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \dots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

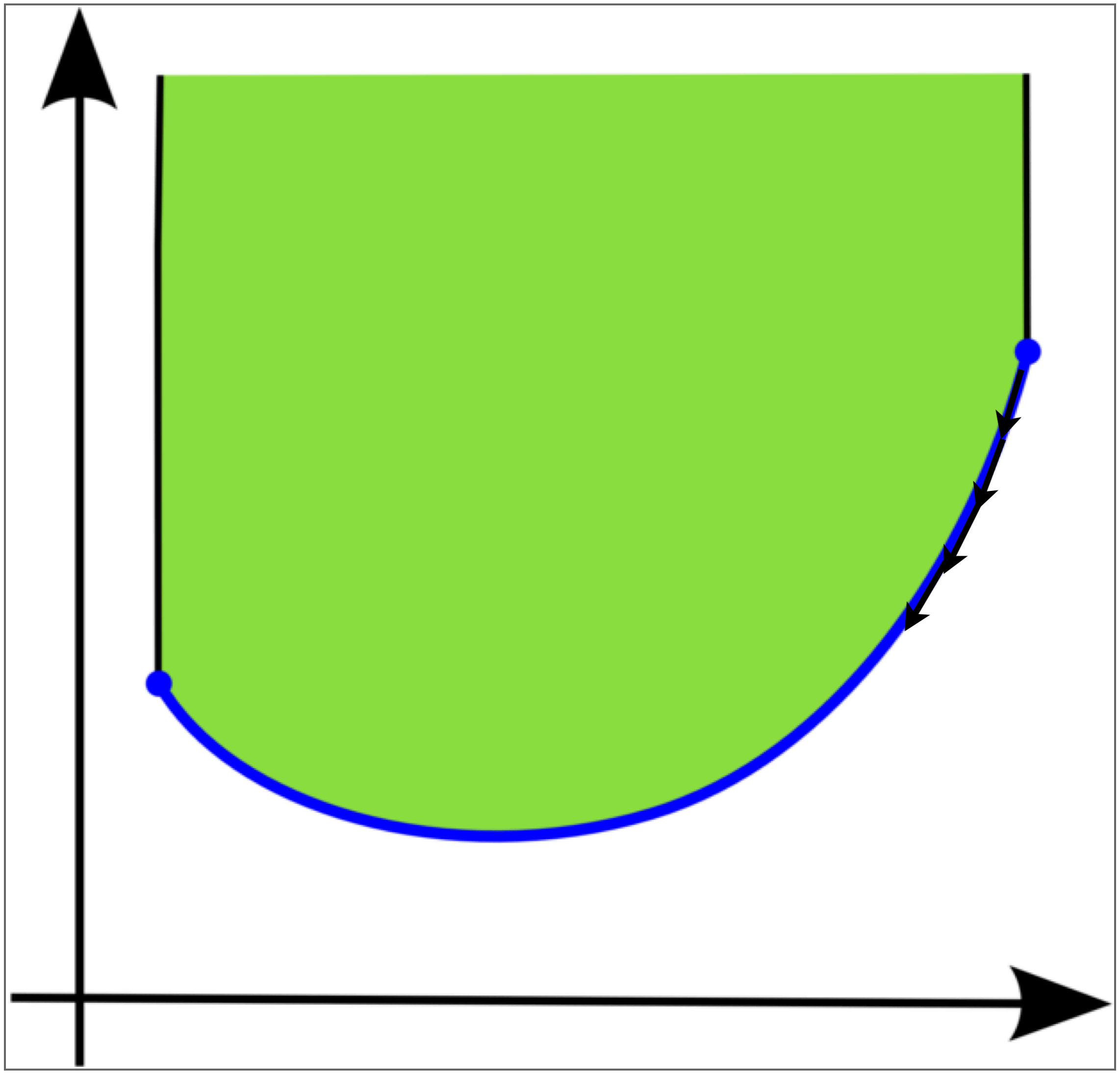
GRADIENT DESCENT

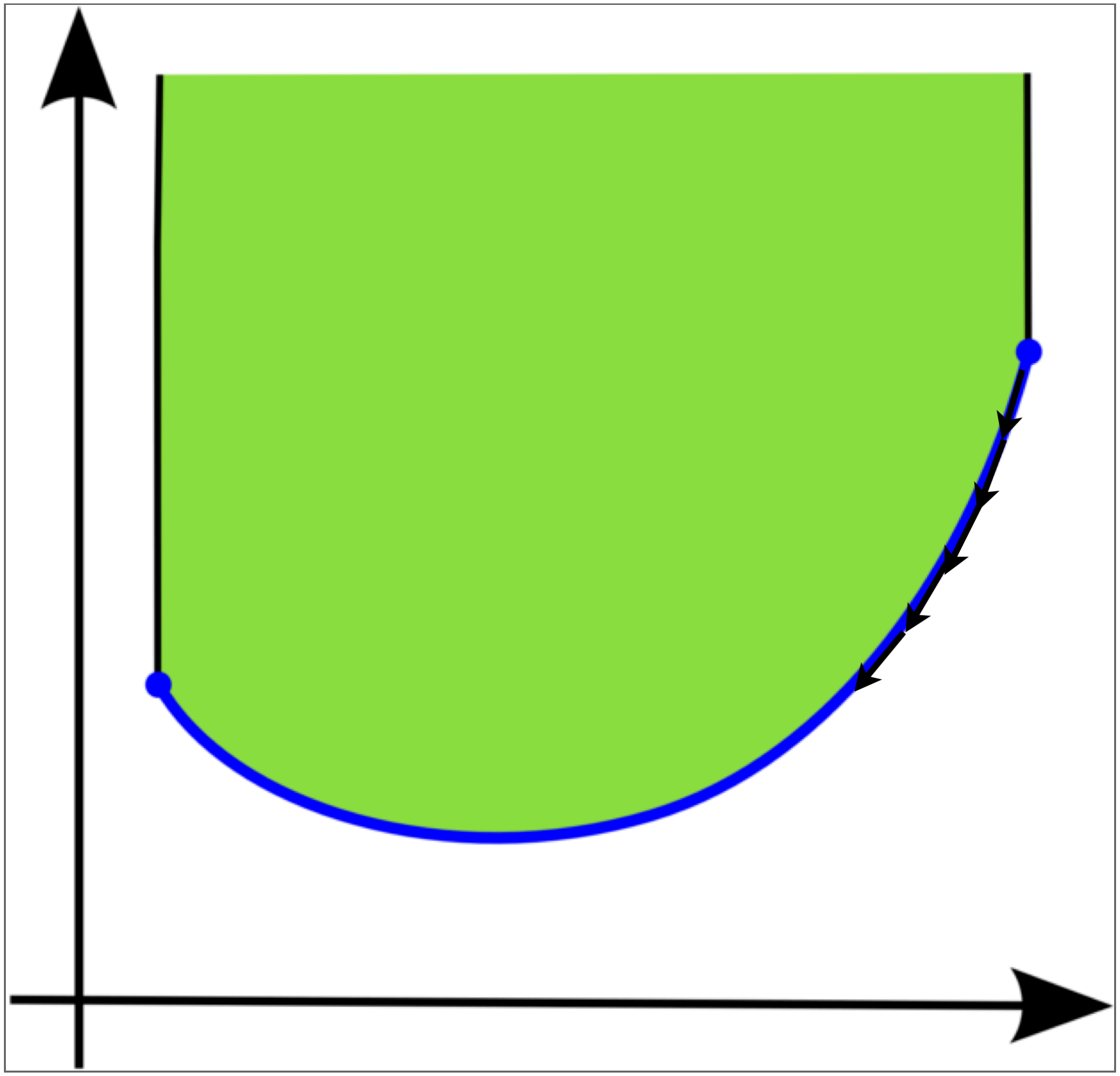


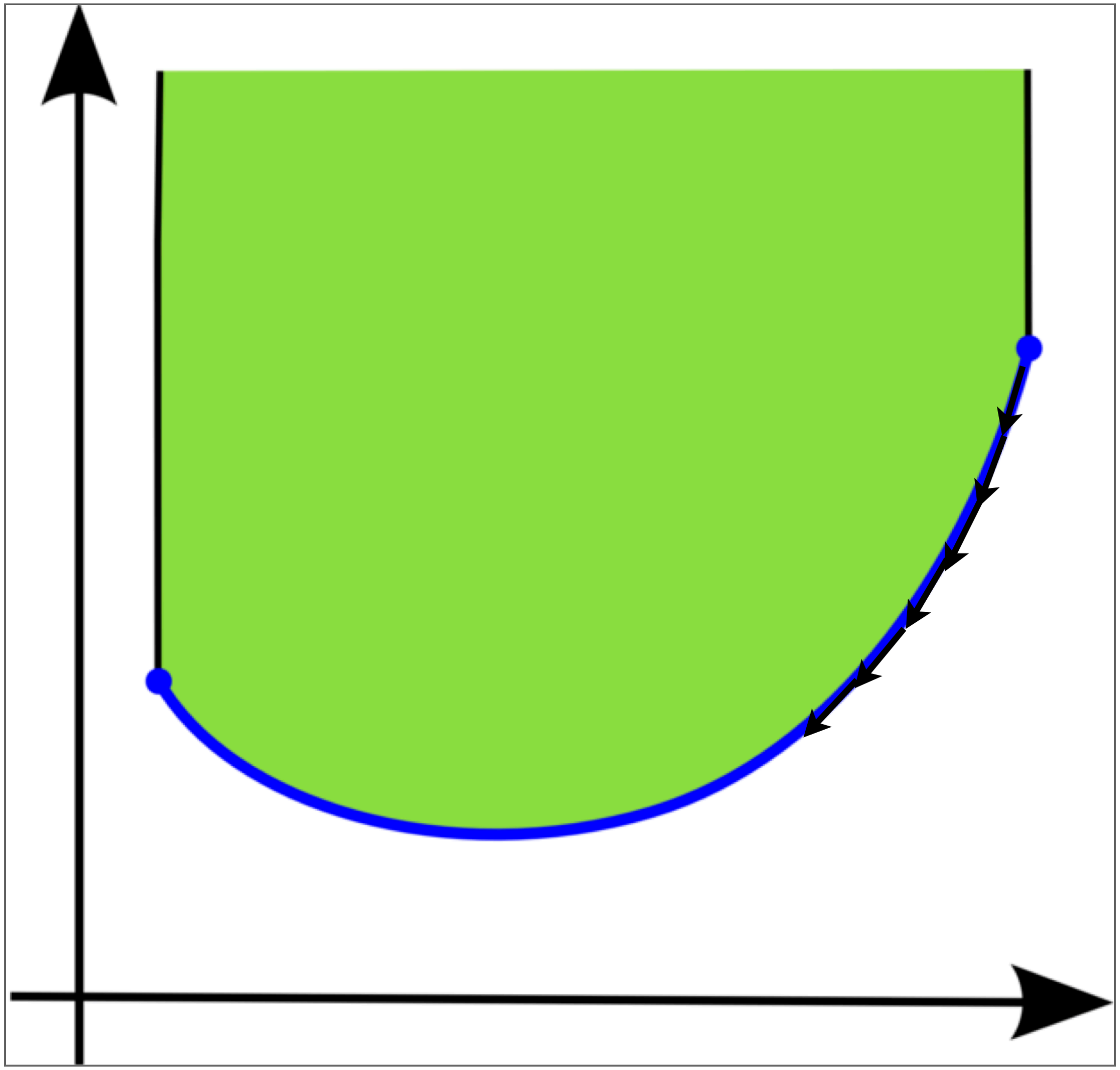












CONVEX OPTIMIZATION: HOW TO SOLVE?

Gradient descent: iteratively update the value of $\mathbf{x}$

- A simple algorithm for unconstrained optimization $\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$

Gradient Descent:

Initialize $\boldsymbol{x} \leftarrow \boldsymbol{x}_0$;
repeat
| $\boldsymbol{x} \leftarrow \boldsymbol{x} - \alpha \bigtriangledown_{\boldsymbol{x}} f(\boldsymbol{x})$;
until *convergence*;

- Variants
- How to choose $\mathbf{x}_0$, e.g., $\mathbf{x}_0 = \mathbf{0}$
- How to choose and update step-size α, e.g., trial and error, line-search methods etc.
- How to define "convergence", e.g., $\|\mathbf{x}^{i+1} - \mathbf{x}^i\| \leq \epsilon$

PROJECTED GRADIENT DESCENT

Iteratively update the value of $\mathbf{x}$ while ensuring $\mathbf{x} \in \mathcal{F}$

Gradient Descent:

Initialize $\boldsymbol{x} \leftarrow \boldsymbol{x}_0$;
repeat
| $\boldsymbol{x} \leftarrow P_{\mathcal{F}}(\boldsymbol{x} - \alpha \bigtriangledown_{\boldsymbol{x}} f(\boldsymbol{x}))$;
until *convergence*;

- $P_{\mathcal{F}}$ projects a point to the constraint set.
- Variant:
- How to choose $P_{\mathcal{F}}$, e.g., $P_{\mathcal{F}} = \underset{\mathbf{x}' \in \mathcal{F}}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{x}'\|_2^2$

CONVEX OPTIMIZATION: HOW TO SOLVE?

Unconstrained and differentiable

- Gradient descent
- Set derivative to be 0
- Closed form solution
- (Not covered) Newton's Method (if twice differentiable)

Constrained and differentiable

- (Not covered) Projected gradient descent
- (Not covered) Interior Point Method

(Not covered) Non-differentiable

- ϵ-Subgradient Method

Cutting Plane Method

CONVEX OPTIMIZATION: APPLY

Model a problem as a convex optimization problem

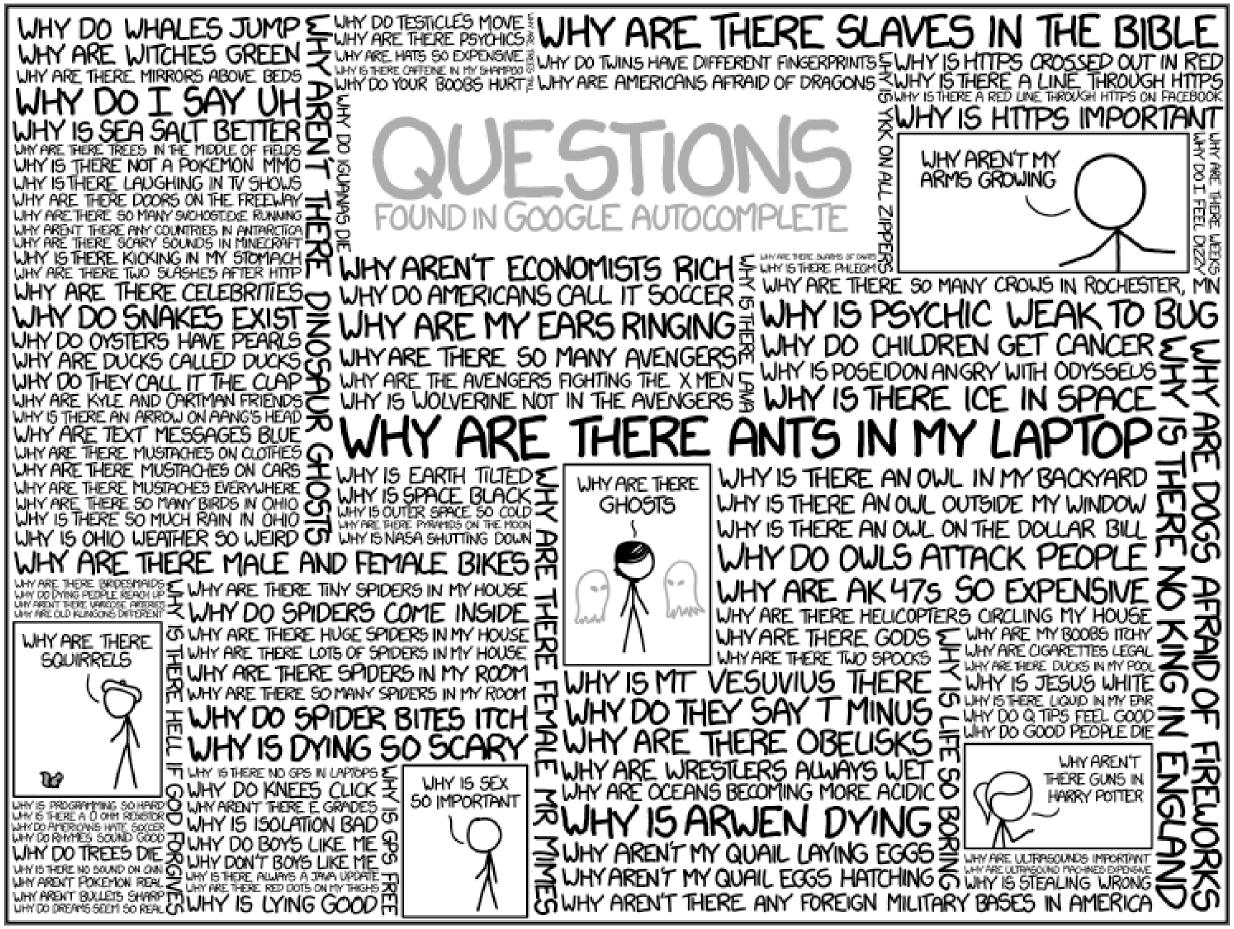
- Define variable, feasible set, objective function
- Prove it is convex (convex function + convex set)

Solve the convex optimization problem

- Build up the model
- Call a solver
- Examples: fmincon (MATLAB), cvxpy (Python), cvxopt (Python), cvx (MATLAB)

Map the solution back to the original problem

Q & A


WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD
WHY AREN'T THERE DINOSAUR GHOSTS
WHY DO IGUANAS DIE
WHY DO TESTICLES MOVE
WHY ARE THERE PSYCHICS
WHY ARE HATS SO EXPENSIVE
WHY IS THERE CAFFEINE IN MY SHAMPOO
WHY DO YOUR BOOBS HURT
WHY ARE THERE SLAVES IN THE BIBLE
WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS
QUESTIONS
FOUND IN GOOGLE AUTOCOMPLETE
WHY IS YKK ON ALL ZIPPERS
WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT
WHY AREN'T MY ARMS GROWING
WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY
WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS
WHY IS THERE LAVA
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE
WHY ARE THERE ANTS IN MY LAPTOP
WHY IS EARTH TILTED
WHY IS SPACE BLACK
WHY IS OUTER SPACE SO COLD
WHY ARE THERE PYRAMIDS ON THE MOON
WHY IS NASA SHUTTING DOWN
WHY ARE THERE GHOSTS
WHY ARE THERE FEMALE MR MIMES
WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK47S SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS
WHY IS THERE NO KING IN ENGLAND
WHY ARE DOGS AFRAID OF FIREWORKS
WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY
WHY ARE THERE SQUIRRELS
WHY IS THERE HELL IF GOD FORGIVES
WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD
WHY IS GPS FREE
WHY IS SEX SO IMPORTANT
WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL
WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA
WHY IS LIFE SO BORING
WHY ARE MY BOOBS ITCHY
WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE
WHY AREN'T THERE GUNS IN HARRY POTTER
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

XKCD

Speaker notes