Starting Sept 26!

- Convex (but non-linear) optimization
- Summary of Project Ideas
- Quadratic Programming! "PSD"
- TSP and Lazy Constraints (Modifying $\left\{ \begin{matrix} LP \\ IP \end{matrix} \right\}$)
- PROPOSAL GRADING (page 17)
- Friday, Oct 5: Getting a good/better model (page 20)

Sept 26:

Most real world problems are not truly linear ...

General problem: max $f(\vec{x})$ s.t. $\left\{ \begin{matrix} \text{constraints} \\ \text{on } \vec{x} \end{matrix} \right\}$

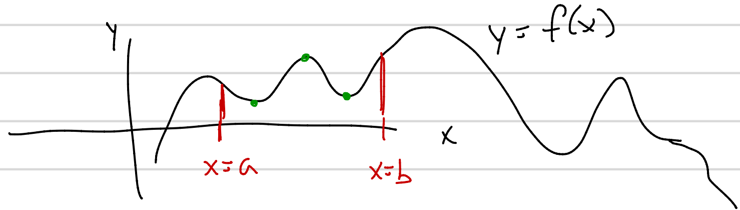
"non-linear program"

"methematical program"

We need some conditions on f and $\left\{ \begin{matrix} \text{constraints} \\ \text{on } \vec{x} \end{matrix} \right\}$ to get something that a computer can solve...

What makes LP work?

Think of calculus optimization:

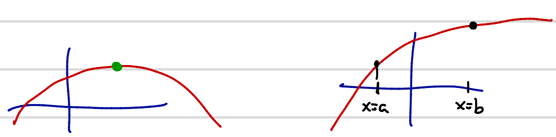


(1) Find x s.t. $f'(x) = 0$: candidates for local max / local min

(2) If constrain x to be $a \le x \le b$, then check $f(a)$, $f(b)$

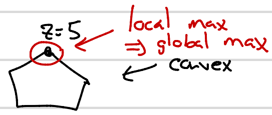
=

If $f(x)$ is concave down, and you want max f:



A local maximum on $\mathbb{R}$ or interval is a global maximum

LP works on this idea.



CP = Convex programming: feasible region convex, f is concave down

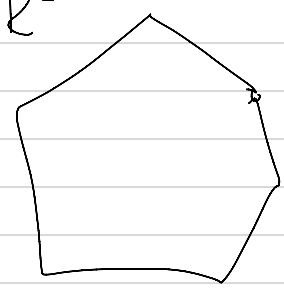
QP = Quadratic programming: CP and f is quadratic function

Quadratic programming:

max $f(\vec{x})$ s.t. ← f is quadratic, but concave-down

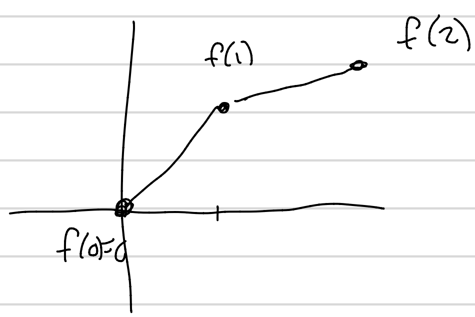
$$\left.\begin{array}{l} A\vec{x} \leq \vec{b} \\ \vec{x} \geq 0 \\ \vdots \end{array}\right\} \text{convex set}$$

in $\mathbb{R}^2$



Examples of quadratic $f(\vec{x})$ that are concave down

(1) $f(\vec{x})$ is a utility with "diminishing returns"



concave down

$x = 0$ utility 0

$x = 1$ " $f(1)$

$x = 2$ " $< 2 \cdot f(1)$

(2) Markowitz model (discuss in a few weeks)

$$f(\vec{x}) = \underbrace{\sum_{i=1}^{n} \alpha_i x_i}_{\text{return}} - \mu \underbrace{\sum_{i,j} x_i x_j \beta_{ij}}_{\text{variance of portfolio}}$$

x_1 = hold of stock 1

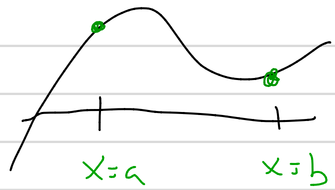
x_2 = " " " 2

$\vdots$

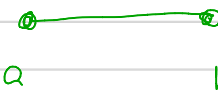
Why is checking the endpoints bad?

- In 1-dim:

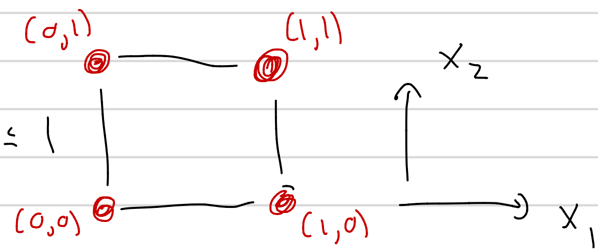
2 endpoints
$x=a, \ x=b$



$a \le x \le b$



- In 2-dim



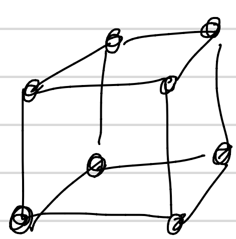
$0 \le x_2 \le 1$

$0 \le x_1 \le 1$

4 "endpoints" (corners)

to check

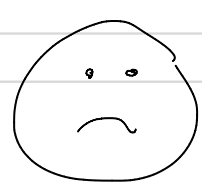
- In 3-dim



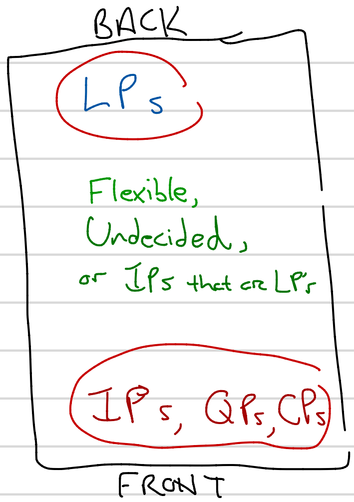
8 endpoints/corners/vertices

- In 20-dim

$2^{20} = 1{,}048{,}576$ vertices



My suggestion :



- Resource Allocation
  - Furniture
  - Diets
- Matrix games

- Weighted Matching
- MOST CONSTRAINED OPIMIZATION

- Bin packing
  - Work allocation, etc.
- Graph Colouring
  - Exam Scheduling
  - Suddku
- Markowitz Model
- Diets that prefer a variety of foods

Typical Challenges:

LP's : Gathering data on variables and constraints

– , ...m. what are interesting constraints

LP's with 100 decision vars constraints have solutions

with most decision variables $= 0$

Sept 28:

- $\underbrace{\text{Blessing}}_{\text{Examples today}}$ / Curse of the simplex method, really LP...

  ☹ for a lot of models

- In a lot of projects, a <u>quadratic</u> objective func is better (more realistic, gives better optimal solutions, etc.)

  However, you often need quadratic to be concave down...

  [Gurobi will refuse to solve it otherwise, ...]

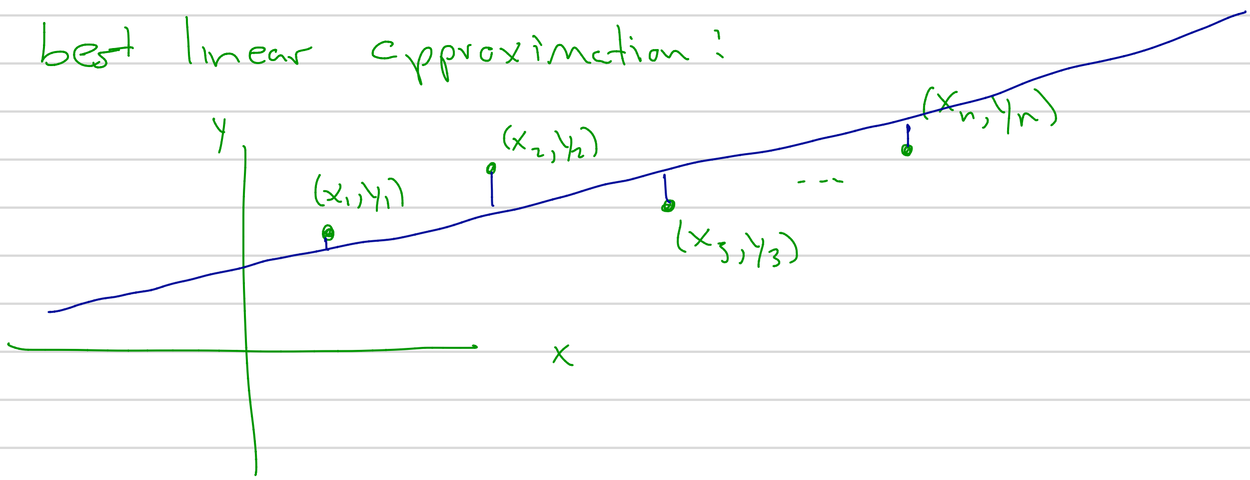
- Many QP's are naturally concave down...

=

Homework #2: Some problems have many IP formulations...

Blessing of simplex method: it produces optimal solutions with some guarenteed # of zeros...

E.g. best linear approximation:



Want $y = ax + b$ that fits the data well.

For each point (x_i, y_i), $\text{error}_i = |y_i - ax_i - b|$

(1) linear regression is: choose a, b s.t. $\sum_{i=1}^{n} (\text{error}_i)^2$ is minimal.

(2) in some situations, want to minimize

$$f(a,b) = \max_i (\text{error}_i)$$

"best max error fit" "best L^∞ fit" "best Chebyshev fit"
Tcheb-

(linear approximations: $y = ax + b$)

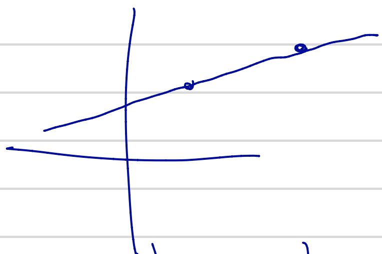
(similar properties for $y = ax^2 + bx + c$
$y = ax^2 + bx + c + de^{-x}$
⋮)

Given $(x_1,y_1),\ldots,(x_n,y_n)$ want a,b s.t.

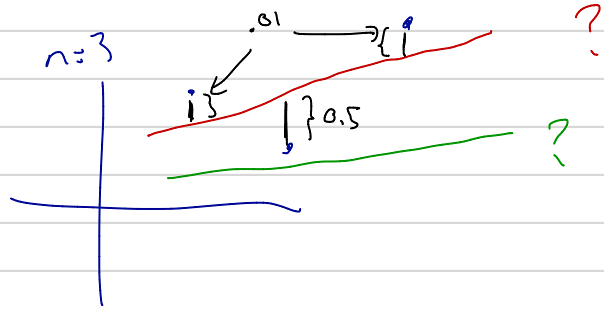
$$\max_{i=1,\ldots,n} |y_i - ax_i - b| \quad \text{is as small as possible.}$$

What do we expect?

$n=2$

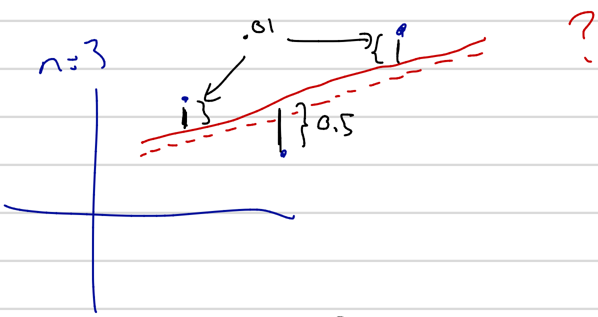


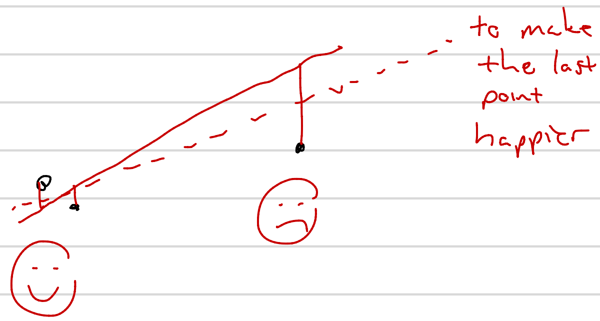
$n=3$

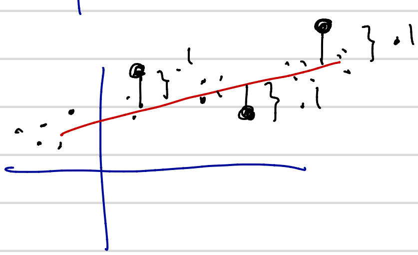


Can always do exactly

$n=3$







A "support" point is an (x_i,y_i) s.t. for the best a,b,

$$|y_i - ax_i - b| \quad \text{is largest}$$

LP: min w s.t.

$$-w \le y_1 - ax_1 - b \le w$$
$$-w \le y_2 - ax_2 - b \le w$$
$$\vdots$$
$$-w \le y_n - ax_n - b \le w$$

Decision vars a,b,w

(Given $x_1,\ldots,x_n$, $y_1,\ldots,y_n$)

$w=0$ means perfect fit

usually we expect $w>0$

[we can assume best $a,b > 0$]

First dictionary:

decision vars a, b, w

slack variables

2n inequalities

2n slack vars

$$u_1 = w - (y_1 - ax_1 - b) \geq 0$$

$$u_2 = (y_1 - ax_1 - b) + w \geq 0$$

$$\vdots$$

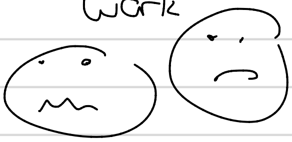
2 slack variables

$$-w \leq y_1 - ax_1 - b \leq w$$

In the initial dictionary

$\underbrace{\text{2n basic vars}}_{\text{slack}}$ $\underbrace{\text{3 non-basic vars}}_{\text{decision}}$

work



optimal solution

$\underbrace{\text{(rest) basic are} \geq 0}$

a, b, w

$\underbrace{\text{3 non-basic vars}}_{\text{3 slack variables}} \to 0$

comes from

$$y_i - ax_i - b \leq w$$

OR

$$-w \leq y_i - ax_i - b$$

So

$$y_i - ax_i - b = w$$

OR

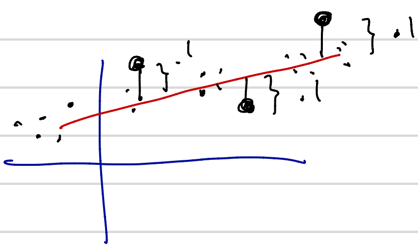
$$-w = y_i - ax_i - b$$

three times

So

$y_i - ax_i - b = w$
OR
$-w = y_i - ax_i - b$

three times



A "support" point is an (x_i, y_i) s.t. for the best a, b,

$|y_i - ax_i - b|$ is largest

simplex method $\Rightarrow$ if 3 decision vars, and they are all $>$ in opt, $\Rightarrow$ 3 slack vars are 0 in optimality

Bad news: if you have many decision vars & few slack vars

Oct 1

- Please email me: jf@math.ubc.ca if you want to work in a group of ≤ 2 people

- Remarks on QP's and what Gurobi (and you) want to avoid

- TSP's: Lazy constraints, "modifying LP"

- Modifying other LP's/IP's $\begin{cases} \text{curve fitting } (L^\infty, L^1 \to L^2) \\ \text{airplane scheduling } (\text{max} \to \text{sum}) \end{cases}$

Quadratic Programs:

$$\text{Objective}(\vec{x}) = \text{Linear}(\vec{x}) + \text{Quadratic}(\vec{x})$$

$\vec{x}$ = decision vars

$$\text{Linear}(\vec{x}) = \vec{c}\cdot\vec{x} = \sum_{i=1}^{n} c_i x_i \qquad (\vec{x} \in \mathbb{R}^n, \mathbb{Z}^n)$$

$$\text{Quadratic}(\vec{x}) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$$

concave-down or ☹

Good new: $Q(\vec{x}) = \sum \alpha_{ij} x_i x_j$ is concave-down iff $Q(\vec{x}) \le 0$ all $\vec{x}$.

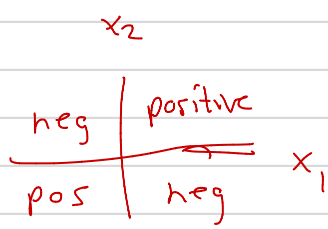
Example:

max $U(\vec{x}) = U(x_1, x_2) = x_1 + x_2 + \underline{\qquad}$

reward you for having some of x_1, some of x_2

☹ example: $x_1 + x_2 + \underbrace{(0.01)\, x_1 x_2}_{\text{is not concave-down}}$

look at $z = x_1 x_2$

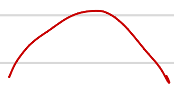


What works

$$\max U(\vec{x}) = x_1 + x_2 - (0.1)\underbrace{(x_1 - x_2)^2}_{\text{concave down}} - (x_1 - 3x_2)^2$$

$$- (4x_1 - x_2)^2$$

$$\vdots$$



$-(0.1)(x_1 - x_2)^2$ penalized for not same of each, not equal amounts of each

$$= \text{Linear}(x_1, x_2) + (ax^2 + bxy + cy^2)$$

Standard tests: given $Q = Q(\vec{x})$ quadratic, is it concave down? i.e. always ≤ 0?

Markowitz model: $Obj = \text{Linear}(\vec{x}) - \mu \underbrace{\text{Variance}(\vec{x})}$

In Markowitz model, you are guarenteed that the quadratic term is always ≤ 0, hence concave down.

Next:

Travelling Salesman Problem   TSP (big field)
person

Idea: n cities

Cost

C_{ij} = going from city i to city j

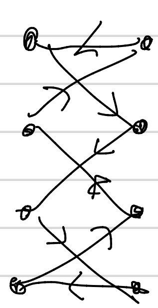
Start at city #1, want to visit all cities with the lowest cost

=

TSP used in a lot of problems...

Toy problem: lacing your shoes

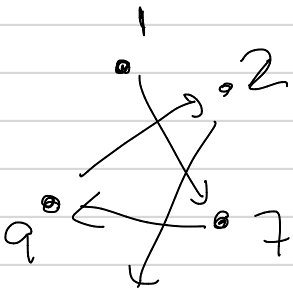
4 eyelets, left and right, want to run lace thru all eyelets



tie shoes

Usual formulation:

$$x_{ij} = \begin{cases} 1 & \text{from city } i \longrightarrow \text{city } j \\ 0 & \end{cases}$$



etc.

What equalities or inequalities?

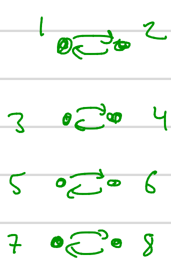
$$\sum_{i,j=1,\dots,n} x_{ij} = n$$

For each i, $\quad \displaystyle\sum_{j=1}^{n} x_{ij} = 1 \quad$ from city i, go to exactly one city afterwards

For each j $\quad \displaystyle\sum_{i=1}^{n} x_{ij} = 1 \quad$ city j comes right after one city

Objective: minimize total cost: $\displaystyle\sum_{i,j=1}^{n} x_{ij} \underbrace{c_{ij}}_{\text{"cost"}}$

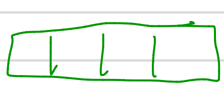
This alone:



$\boxed{\text{2 ways (many ways) to avoid getting more than one loop...}}$

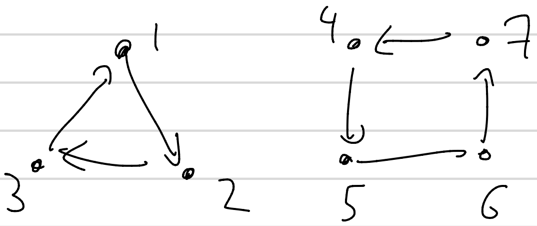
Problem: You can get a set of loops, not just one...

Recall: Sudoku



2 ways of writing IP, one is better than the other

How to prohibit 2 or more loops



There's no connection between $\{1,2,3\}$ and $\{4,5,6,7\}$

$$\sum_{\substack{i=1,2,3 \\ j=4,5,6,7}} x_{ij} \geq 1$$

this is a standard way of solving TSP...

The problem: really

for every $S \subset \{1,\dots,n\}$ $\quad \sum_{i \in S,\ j \notin S} x_{ij} \geq 1$

$\#$ of S roughly 2^n.

Typically run "lazy constraint addition"

Keep running the IP, and prohibit any loops that comes up by adding one of

Rem: Gurobi has this strategy on its website

Oct 3:

Proposals due Oct 5 (Friday), meaning

- 11:59pm on Monday, Oct 8 (= Thanksgiving) on Canvas
    - You can choose your own group under Projects, submit
- 300-700 words, excluding math symbols and formulas (for the Proposal part, not the outline)
- Should have
    - ☆ 1-2 sentence overview
    - ☆ 3 or more questions to investigate
        - ≥1 parametric
        - ≥1 easy
        - ≥1 difficult
    - ☆ At least one precise LP/IP/QP/etc.
    - ☆ List of who is doing what
    - ☆ Should be concise, well-organized, and clear

(One point each)

- Can be resubmitted once, within a week of being returned, for a higher score in JCCS, room X561

Oct 3: Last appointments this week: Th 2-3, Fri 9:30-11:30, Canvas Sign Up

- Gurobi examples of good/bad QP
- More TSP
- { Synthetic Data (for TSP or Flight Schedules or Cities in Canada or ...)
  Toy examples
- Next topics: GOOD, "MEANINGLESS" objective functions

- TSP
  - Many competing formulations
  - Triangle inequality in realistic models
  - Two extreme toy examples
  - Synthetic Data

$\mathrm{dist}(P,R) \leq \mathrm{dist}(P,Q) + \mathrm{dist}(Q,R)$

TSP: $\quad i \xrightarrow{x_{ij}=1} j$

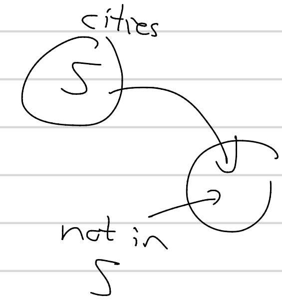
$$\sum_{j=1}^{n} x_{ij} = 1$$

$$\sum_{i=1}^{n} x_{ij} = 1$$

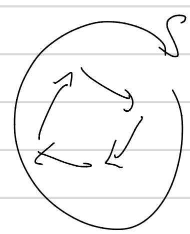
Added

(1) $S \subset \{1,\dots,n\}$, n cities

$$\sum_{i\in S,\, j\notin S} x_{ij} \geq 1$$



(2) For any $S \subset \{1,\dots,n\}$ don't want



$$\sum_{i,j\in S} x_{ij} \leq \underbrace{|S|}_{\text{size of } S} - 1 \qquad (\text{if } S \neq \{1,\dots,n\})$$

(3) etc.

(4)

(5)

⋮

Friday Oct 5:

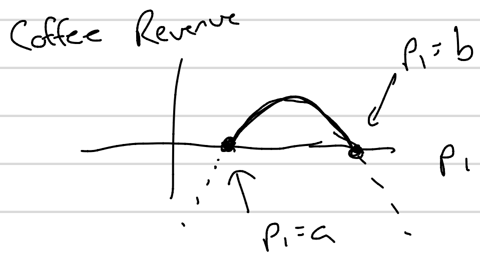
- How to make a better model: Small adjustments

  Based on student projects

1. Optimal revenue: one variable $\rightsquigarrow$ many variables

2. Resource Allocation LP example [diet problem]: find a better parameter

Example of minor adjustment:



decide on price

Selling coffee $\longleftrightarrow$ p_1

Selling donuts $\longleftrightarrow$ p_2

Coffee

$$a \le p_1 \le b$$

$p_1 = a$ means: cost to {cafe / donut shop} is a

$p_1 = b$ means: no one buys if price $> b$

make over cost

Simplest approximation (reasonable): Coffee Rev $= C(p_1 - a)\underbrace{(b - p_1)}_{\text{demand at } p_1}$

$= \text{const}_0 + \text{const}_1 p_1 - C p_1^2, \quad C > 0$

Selling Donuts at p_2

$= C_0 + C_1 p_1 + C_2 p_1^2 \qquad C_2 = -C < 0$

Revenue Donuts $=$ $\longrightarrow$ $= D_0 + D_1 p_2 + D_2 p_2^2$

Total Rev $(p_1, p_2) = (C_0 + D_0) + (C_1 p_1 + D_1 p_2) + \underbrace{\text{Quadratic}}_{C_2 p_1^2 + D_2 p_2^2}$

Solution is not so interesting ☹

Concave Down

☺ Small adjustment:

Rev Coffee: $C\,(p_1 - a)\,(b - p_1 + \underbrace{0.1\,(\$3.99 - p_2)}_{\text{minor adjustment}})$

$\hookrightarrow$ get $-\left(\begin{smallmatrix}\text{small}\\ \text{const}\end{smallmatrix}\right) p_1 p_2$ term

☺ Quadratic $-p_1^2 - 3p_2^2$

to $-p_1^2 - 3p_2^2 + (\text{small})\, p_1 p_2$

if small enough, the quadratic is still concave-down ...

Coffee, Donuts can repel or attract

Diet problem: min cost: $c_1x_1+c_2x_2+\dots+c_{100}x_{100}$

Decision Vars: $x_1,\dots,x_{100}$ ← foods

Constraints:

$$\left\{\begin{array}{l} \text{min protein} \le p_1x_1+\dots+p_{100}x_{100} \le \text{max protein} \\ \text{min carbo} \le s_1x_1+\dots+s_{100}x_{100} \le \text{max carbs} \\ \vdots \end{array}\right.$$

3,4 constraints, bounded above & below

Yields diets with very few types of food consumed.

Could bound each food $x_1 \le m_1$, $x_2 \le m_2, \dots$

Have $m_1,\dots,m_{100}$ or $p_1,\dots,p_{100}$ or $s_1,\dots,s_{100}$ be parameters

What if

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{100} \end{bmatrix} = \underset{\text{drought}}{\alpha} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{100} \end{bmatrix} + \underset{\text{currency parameter}}{\beta} \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_{100} \end{bmatrix} + \begin{bmatrix} \hat{c}_1 \\ \vdots \\ \tilde{c}_{100} \end{bmatrix}$$

2 parameters

constants

current prices

LP with α, β parameters