Lecture #22: Quadratic Programming

Today: Important class of optimization problems, with many connections to linear algebra

Let's start with a concrete example:

$$\min_{x,y} \; 2x^2 - 2xy + y^2 - 2\sqrt{2}\,x + 4\sqrt{2}\,y$$

Q: what is the optimal solution?

This is an <u>unconstrained</u> optimization problem

Later we will allow <u>constraints</u>, that the variables have to belong to some region

I claim that eigendecomp. can help us!

Step # 1: Write the QP in matrix-vector notation:

$$\min_{x,y} \; [x \; y] \begin{bmatrix} 2 & -2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [-2\sqrt{2} \; +4\sqrt{2}] \begin{bmatrix} x \\ y \end{bmatrix}$$

Actually there is a more convenient way to write this

$$\min_{x,y} \; [x \quad y] \underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{z} + \underbrace{[-2\sqrt{2} \; +4\sqrt{2}]}_{b^T} \begin{bmatrix} x \\ y \end{bmatrix}$$

Now A is symmetric, and we know a lot about existence/structure of eigen decomp.

We will write things more compactly as:

$$\min_{z} \; z^T A z + b^T z$$

Let's see if using an eigen decomp. simplifies things

$$A = UDU^T$$

Q2: what do we know about U?

It is orthogonal

what property of A ensures this?

A is symmetric

Concretely, we find:

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

This tells us a convenient <u>change of variables:</u>

Let $$\underbrace{\begin{bmatrix} x' \\ y' \end{bmatrix}}_{z'} = \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{U^T} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{z}$$

Now let's see what this does

$$\min_{z} \; z^T U D U^T z \quad + b^T z$$

$$= \quad \min_{z} \; \underbrace{z^T U}_{z'^T} D \underbrace{U^T z}_{z'} \quad + \underbrace{b^T U}_{b'^T} \underbrace{U^T z}_{z'}$$

$$= \quad \min_{x', y'} \; 3x'^2 + y'^2 \quad + 6x' \; + 2y'$$

Does this look simpler? Not quite there yet

Now lets complete the squares:

$$= \min_{x', y'} \; 3(x'+1)^2 + (y'+1)^2 \; - 4$$

Aha! Now it's easy to find optima

$$x' = -1 \, , \; y' = -1 \Rightarrow \text{obj. value } -4$$

Q3: what x, y achieve this in the original problem?

$$\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{U} \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

So what did we learn?

unconstrained QP = eigen decomp. + complete the square

Actually we got lucky; what would happen if, say

$$A = U \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} U^T$$

<u>Poll</u> what is $\min z^T A z + b^T z$?

(a) 0 (b) -4 (c) $-\infty$ (d) $+\infty$

<u>Hint:</u> Think about it using the <u>same</u> change of variables we did before

Again, consider the optimization problem with x', y':

$$\min_{x', y'} [x' \ y'] \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \text{stuff}$$

$$\xRightarrow[\text{the square}]{\text{complete}} \quad 3(x' + a_1)^2 - 1(y' + a_2)^2 + C$$

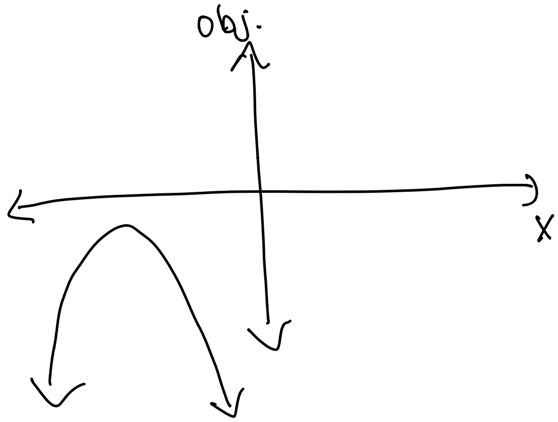
Now what is the optima? $-\infty$

<u>Lemma:</u> In unconstrained QP

(1) If symmetric A has a negative eigenvalue then optimum is $-\infty$

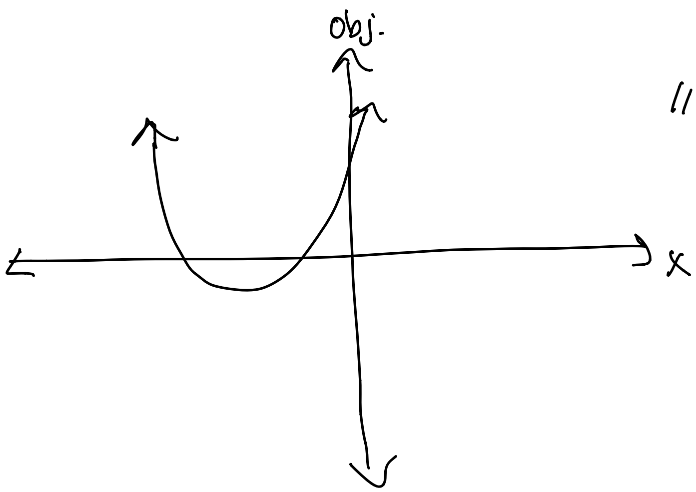
(2) If A has nonnegative eigenvalues (called <u>positive semidefinite</u> PSD) then optimum is finite

There's a geometric picture that explains this lemma what does (1) look like?



"concave"

what does (2) look like?



"convex"

In the setting of optimization:

"convex" = tractable

"concave" = probably not

Actually, the eigenvalues tell us even more:

Q4: Is the solution unique?

Poll: For unconstrained QP with PSD A, is there a unique solution?

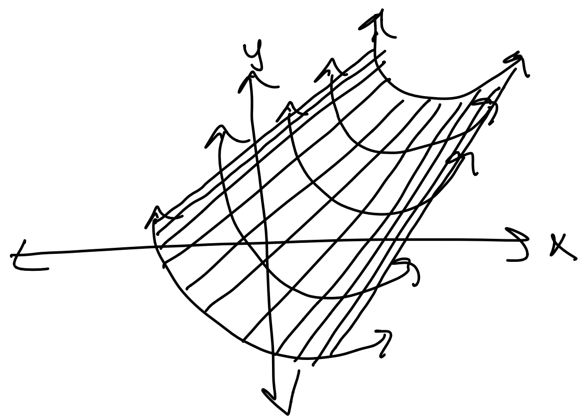
(a) Yes (b) No, never (c) Sometimes

Lemma: In unconstrained QP with PSD A we have:

the optimal choice of z is unique iff A has only positive eigenvalues

Again, there is a geometric picture. Suppose A is PSD but not PD

positive def.



Or, to put it another way, if there is a $c \in N(A)$ you get a family of optima

$$z + \underset{\substack{\uparrow \\ \text{scalar}}}{\alpha} c$$

Now let's add constraints:

$$\min \frac{x^T P x}{2} + q^T x$$
$$\text{s.t. } Ax = b$$

equality constrained QP

This is now a very expressive class of problems:

Lemma: Least squares can be written as an equality constrained QP

Recall in the underdetermined case:

$$\text{(LS)} \quad \min \|x\|^2 \quad \text{s.t. } Ax = b$$

← pick out "simplest"

How do we write this as constrained QP?

$$\|x\|^2 = x^T I x$$

Now let's revisit an application from Lecture #3 from a new perspective

Along the way let's answer a question that might be on your mind:

We know Julia can solve problems on your psets, but what about nature?

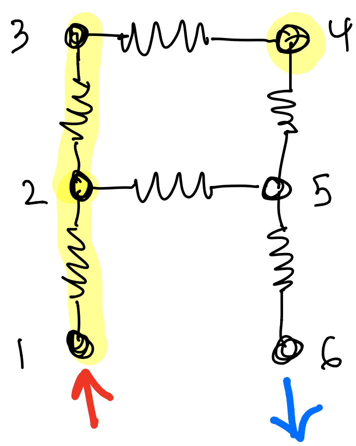
Or to put it another way:

Is nature able to solve interesting optimization problems?

<u>Lemma:</u> In an electrical circuit the actual current solves an <u>energy min. prob.</u>

least squares

In Lecture #3 we talked about this circuit:



1 unit current in/out

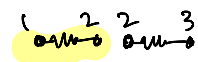
We set up a matrix that is called the <u>Laplacian</u>

$$L = \begin{bmatrix} 1 & -1 & \\ -1 & 1 & 0 \\ & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & & & 0 \\ & 1 & -1 & \\ & -1 & 1 & \\ 0 & & & 0 \end{bmatrix} + \dots$$

Each matrix in the sum represents a resistor

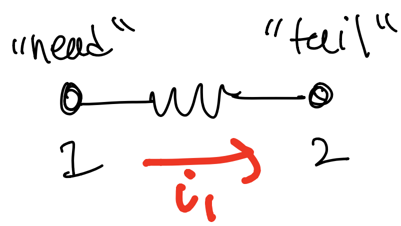
<u>Key Point:</u> There's another way to write the Laplacian:

$$L = \overset{n \times m}{A} A^T$$



$$A = \begin{bmatrix} 1 & 0 & \cdots \\ -1 & 1 & \\ 0 & -1 & \\ \vdots & \vdots & \end{bmatrix}$$

In particular choose an arbitrary <u>orientation</u> of each resistor



and put +1 in head and −1 in tail

Now let's express the energy minimization problem as least squares:

$$\min \; \|i\|^2$$
$$\text{s.t.} \quad A\, i = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{bmatrix}$$

Recall from Lecture #12 the optimal solution is:

$$i^* = A^T \underbrace{(A A^T)}_{L}{}^{-1} b$$

Hence we have

$$\underset{\text{voltages}}{v} = L^{-1} b \implies A^T v \text{ the currents}$$

Whoa! How the current chooses to split comes from least squares/QP

<u>Next Time:</u> Does nature always find the optimal solution?

Alternatively: when have we reached a "local" vs. "global" optimum, and how do we tell the difference?