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 unconstrained optimization problem

Later we will allow constraints, that the variables have to belong to some region

I claim that eigen decomp. can help us!
Step #1: Write the QP in matrix-vector notation:

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

Actually there is a more convenient way to write this:

$$\min_{x,y} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -2\sqrt{2} + 4\sqrt{2} \end{bmatrix} \begin{bmatrix} y \end{bmatrix}$$

$$A \quad z = b^T$$

Now $A$ is symmetric, and we know a lot about existence (structure of eigen decomps.)

We will write things more compactly as:

$$\min_z \begin{bmatrix} z^T A z + b^T z \end{bmatrix}$$
Let's see if using an eigen decompostion simplifies things

$$A = U D U^T$$

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 change of variables:

Let $$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$x', y'$$
Now let's see what this does

$$\min \ z^T U D U^T z + b^T z$$

$$= \min \ z^T U D U^T z + b^T U U^T z$$

$$= \min \ 3x'^2 + y'^2 + 6x' + 2y'$$

Does this look simpler? Not quite there yet.

Now let's complete the squares:

$$= \min \ 3(x' + 1)^2 + (y' + 1)^2 - 4$$

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?

\[
\begin{bmatrix}
  x \\
  y
\end{bmatrix} = \begin{bmatrix}
  -\frac{1}{2} & \frac{1}{2} \\
  \frac{1}{2} & \frac{1}{2}
\end{bmatrix}
\begin{bmatrix}
  -1 \\
  u
\end{bmatrix}
\]

So what did we learn?

unconstrained = eigen decom. + complete the square

Actually we got lucky; what would happen if, say

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

Poll what is min \( z^T A z + b^T z \)?

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

Hint: Think about it using the same change of variables we did before.
Again, consider the optimization problem with $x', y'$:

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

complete the square

$$
3(x' + a_1)^2 - 1(y' + a_2)^2 + C
$$

Now what is the optimum? $-\infty$

Lemma: In unconstrained OP

1. If $A$ has a negative eigenvalue, then optimum is $-\infty$

2. If $A$ has non-negative eigenvalues (called positive semidefinite, PSD), then optimum is finite

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

What does (2) look like?

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 $c \in \mathcal{N}(A)$ you get a family of optima

\[ z + \alpha c \]

^ scalar

Now let's add constraints:

\[ \min \quad \frac{x^T P x}{2} + c^T x \]

\[ \text{st.} \quad A x = b \]

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:

\[(LS) \quad \min ||x||^2 \Rightarrow \text{pick out "simplest"}
\]

st. \(Ax = b\)

How do we write this as constrained OP?

\[||x||^2 = x^T \Sigma x\]

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?

Lemma: In an electrical circuit the actual current solves an energy min prob. least squares

In Lecture #3 we talked about this circuit:

We set up a matrix that is called the Laplacian

\[
L = \begin{bmatrix}
1 & -1 & 0 \\
-1 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix} + \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & -1 \\
0 & -1 & 0
\end{bmatrix} + \ldots
\]
Each matrix in the sum represents a resistor.

Key point: There's another way to write the Laplacian:

\[
L = A A^T
\]

\[
A = \begin{bmatrix}
1 & 0 & \cdots \\
-1 & 1 & \cdots \\
0 & -1 & \ddots \\
\vdots & \ddots & \ddots
\end{bmatrix}
\]

In particular, choose an arbitrary orientation of each resistor:

```
    "head"     "tail"
    \[ \begin{array}{c}
          \[ \end{array} \\
    \end{array}
\]
```

and put +1 in head and -1 in tail.

Now let's express the energy minimization problem as least squares:
\[
\begin{align*}
\min & \quad \| \mathbf{i} \|^2 \\
\text{st.} & \quad \mathbf{A} \mathbf{i} = \begin{bmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} 
\end{align*}
\]

Recall from lecture #12 the optimal solution is:
\[
\mathbf{i}^* = \mathbf{A}^\top (\mathbf{A} \mathbf{A}^\top)^{-1} \mathbf{b} 
\]

Hence we have:
\[
\mathbf{v} = \mathbf{L}^{-1} \mathbf{b} \implies \mathbf{A}^\top \mathbf{v} \text{ the currents,}
\]
\[
\mathbf{v} \text{ the voltages}
\]

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

**Next Time:** Does nature always find the optimal solution?

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