MECH 6327 - Homework 1

Jonas Wagner

2021, February 3

1 Optimization Problem Formulation

Find and research an example (ideally from your own research experience/interests) of an optimization problem in the real world. Write a short narrative of around 300 words to describe the optimization variables, objective function(s), constraint function(s), and formulate the problem as a mathematical optimization problem. Briefly discuss your knowledge of how computationally easy or difficult the problem is, the number of variables and constraints in typical problem instances, available algorithms or software for solving the problem, etc. (need not be formal, this is just to assess and get you to think about the current state of your knowledge about the problem).

Solution:
When dealing with ensuring the security of a cyber-physical systeman important component is the ability to detect if an attacker is occurring while also minimizing the number of false positives. Often what is done to detect if any sensors have been compromised is to compare the predicted state estimate with the sensor reading. In this case an optimization problem can be conducted to develop a detector that maximizes the detection rate for a given false positive value. Additional constraints on the detector can also be added, such as restricting system state estimates or adding a computational or refresh rate restriction. Additionally, an entirely different cost function could be added to find an optimal solution to minimize or maximize properties. Untimely in this problem more constraints are necessary to find the optimal detector. [1]

2 Math Review

2.1 Cauchy-Schwarz Inequality Proof

Problem: Prove the Cauchy-Schwarz inequality, which states that $\forall x, y \in \Re^n$, there holds

$$|\langle x|y \rangle| \le \|x\|\|y\|,$$

where $\|\cdot\| = \sqrt{\langle \cdot|\cdot \rangle}$.

Solution: Let $\|x\| = \|y\| = 1$.
Consider

$$\|x - y\|^2 = \left(\sqrt{\langle x - y|x - y \rangle} \right)^2 = \langle x - y|x - y \rangle$$

$$\langle x - y|x - y \rangle = \langle x|x - y \rangle + \langle -y|x - y \rangle$$

$$= \langle x|x \rangle + \langle x|-y \rangle + \langle -y|x \rangle + \langle -y|-y \rangle$$

$$= \|x\|^2 - \langle x|y \rangle - \langle y|x \rangle + \|y\|^2$$

$$\|x - y\|^2 = \|x\|^2 - 2\langle x|y \rangle + \|y\|^2$$

$$2\langle x|y \rangle = \|x\|^2 + \|y\|^2 - \|x - y\|^2$$

Now taking the assumption of unit vectors from before and substituting,

$$2\langle x|y \rangle = 1^2 + 1^2 - \|x - y\|^2$$

$$2\langle x|y \rangle = 2 - \|x - y\|^2$$

$$\langle x|y \rangle = 1 - \frac{\|x - y\|^2}{2}$$

Since $\|x\|\|y\| = 1$,

$$\langle x|y \rangle = \|x\|\|y\| - \frac{\|x - y\|^2}{2}$$

Thus, for all unit vectors,

$$\langle x|y \rangle \le \|x\|\|y\|$$

This demonstration can be expanded to include any vectors becouse all vectors can be decomposed into a unit vector and a scaler magnitude:

$$\mathbf{z} = \frac{\hat{\mathbf{e}}_z}{\|z\|}$$

and both the inner product and norms are homogenous with respect to scaler quantities.

2.2 Norm Duality

Problem: Prove the dual norm relationship for the 1-, 2-, and ∞-norms.

2.2.1 1- and ∞-norm

$$\|z\|_{1*} = \max z^T x \mid \|x\|_1 \le 1$$
$$\max z^T x = |\langle z|x\rangle| \le \|z\|_1 \|x\|_1$$
$$= \sum_{i=1}^{n} z_i x_i \le \left(\sum_{i=1}^{n} |z_i|\right)(1)$$
$$= \frac{z^T x}{\|z\|_1} \le 1$$

Since

$$\frac{z^T x}{\|z\|_1} = 1$$

occurs when x is chosen to isolate the maximum value, i.e. $z = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\max z^T x = \max z_i = \|z\|_\infty$$

2.2.2 ∞- and 1-norm

$$\|z\|_{\infty*} = \max z^T x \mid \|x\|_\infty \le 1$$
$$\max z^T x = |\langle z|x\rangle| \le \|z\|_\infty \|x\|_\infty$$
$$= \sum_{i=1}^{n} z_i x_i \le \left(\max_i z_i\right)(1)$$
$$= \sum_{i=1}^{n} \frac{z_i x_i}{\|z\|_\infty} \le 1$$

Since

$$\frac{z^T x}{\|z\|_\infty} = 1$$

occurs when x maximizes the sum, i.e. $z = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $x = \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $z^T x = 5$

$$\max z^T x = \sum_{i=1}^{n} z_i = \|z\|_1$$

3

2.2.3   2- and 2-norm

$$\|z\|_{2*} = \max z^T x \mid \|x\|_2 \leq 1$$

$$\max z^T x = \langle z|x \rangle \leq \|z\|_2 \|x\|_2$$

$$= \sum_{i=1}^{n} z_i x_i \leq \left(\sqrt{\sum_{i=0}^{n} z_i^2} \right) (1)$$

$$= \sum_{i=1}^{n} \frac{z_i x_i}{\sqrt{\sum_{i=0}^{n} z_i^2}} \leq 1$$

Since

$$\frac{z^T x}{\|z\|_2} = 1$$

occurs when all elements of $x = \dfrac{1}{\|z\|_2}$, thus

$$\max z^T x = \sqrt{\sum_{i=1}^{n} z_i^2} = \|z\|_2$$

2.3 Ploting the Unit Norm Balls

Problem: Plot the unit norm balls $\{x \in \Re^2 \mid \|x\|_p \leq 1\}$ for $p = 1, 2, \infty$.
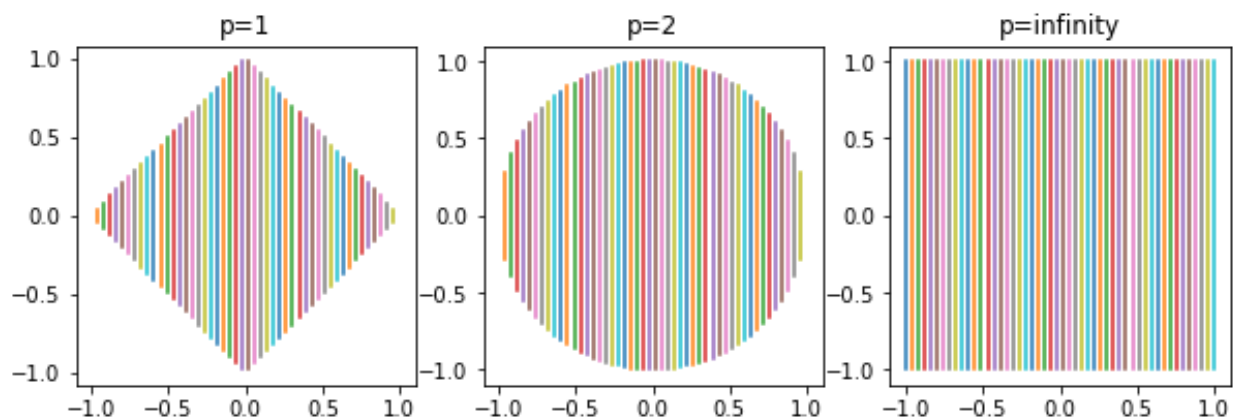
Solution:

Figure 1: Unit Norm Ball Plots

```python
import numpy as np
import matplotlib.plot as plt

x = np.linspace(-1,1)
y1 = 1 - abs(x)
y2 = np.sqrt(1 - x**2)
y3 = 1

fig,axs = plt.subplots(1,3,sharex=True,figsize=(10,10))

axs[0].plot([x,x],[y1,-y1])
axs[0].set_box_aspect(1)
axs[0].set_title('p=1')

axs[1].plot([x,x],[y2,-y2])
axs[1].set_box_aspect(1)
axs[1].set_title('p=2')

axs[2].plot([x,x],[y3,-y3])
axs[2].set_box_aspect(1)
axs[2].set_title('p=infinity')
```

2.4 Eigenvalues Proof 1

Problem: Show that the eigenvalues of $A \in \mathbf{S}^n$ are real.

Solution: Every eigen-value, λ, of matrix, $A \in \mathbf{S}^{n \times n}$, must satisfy the following:

$$Ax = \lambda x, \ \forall x \in \Re^n$$

The following must also be true:

$$x^* A x = x^* \lambda x$$
$$(x^* A x)^* = (x^* \lambda x)^*$$
$$x^* A^* x = x^* \lambda^* x$$

Since, by definition, $A = A^*$,

$$x^* A x = x^* \lambda x = x^* \lambda^* x = x^* A^* x$$
$$\lambda = \lambda^*$$

This means λ must contain no imaginary components and therefore $\lambda \in \Re$.

2.5   Eigenvalues Proof 2

Problem: Show that for $A \in \mathbf{S}^n$,

$$A \succeq 0 \iff \lambda_i\{A\} \geq 0, \ \forall i = 1, \ldots, n$$

Solution: Every eigen-value, $\lambda \in \Re$, of matrix, $A \in \mathbf{S}^{n \times n}$, must satisfy the following:

$$Ax = \lambda x, \ \forall x \in \Re^n$$

By definition, a semi-positive definite matrix must satisfy the following:

$$x^T A x \geq 0, \ \forall x \in \Re^n$$

.

The following must also be true:

$$x^T A x \geq 0$$
$$x^T (\lambda x) \geq 0$$

Thus,

$$\lambda x^T x = x^T A x \geq 0$$

Since $\lambda x^T x \geq 0 \iff \lambda \geq 0$, the following is also true:

$$A \succeq 0 \iff \lambda \geq 0$$

.

2.6 Optimal Control Problem Derivation

Problem: Derive the solution to the finite horizon open-loop optimal control problem:

$$\text{minimize} \sum_{t=0}^{N-1} \left(\|z_t\|_2^2 + \|u_t\|_2^2 \right) + \|z_N\|_2^2$$

where $z_{t+1} = Az_t + Bu_t$ with given z_0. Take the optimization variables to be $z_t \in \Re^n$ for $t = 1, \ldots, N$ and $u_t \in \Re^m$ for $t = 0, \ldots, N-1$.

Solution:

Define the following vectors:

$$\mathbf{z} = \begin{bmatrix} z_1^T \\ \vdots \\ z_N^T \end{bmatrix} \qquad\qquad \mathbf{u} = \begin{bmatrix} u_0^T \\ \vdots \\ u_{N-1}^T \end{bmatrix}$$

$\mathbf{z}$ can then be redefined in terms of $\mathbf{u}$ and z_0:

Since

$$z_{t+1} = Az_t + Bu_t$$
$$z_1 = Az_0 + Bu_0$$
$$z_2 = Az_1 + Bu_1$$
$$\quad = A(Az_0 + Bu_0) + Bu1$$
$$\quad = A^2 z_0 + ABu_0 + Bu_1$$
$$z_3 = A^3 z_0 + A^2 Bu_0 + ABu_1 + Bu_2$$
$$\vdots$$
$$z_N = Az_{N-1} + Bu_{N-1}$$
$$\quad = A^n z_0 + A^{n-1} Bu_0 + \cdots + ABu_{n-1} + Bu_n$$

Additonally, this can be generalized as:

$$z_t = A^t z_0 + \sum_{i=0}^{t-1} A^{t-i} B\mathbf{u}_i$$

The entire sum can be vectorized as follows:

$$\mathbf{z} = \begin{bmatrix} A & | & B & 0 & \ldots & 0 \\ A^2 & | & AB & B & \ldots & 0 \\ \vdots & | & \vdots & \ddots & \ddots & \vdots \\ A^N & | & A^{n-1}B & \ldots & AB & B \end{bmatrix} \begin{bmatrix} z_0 \\ -- \\ \mathbf{u} \end{bmatrix}$$

$$= [A_m | B_m] \begin{bmatrix} z_0 \\ -- \\ \mathbf{u} \end{bmatrix}$$

$$= A_m z_0 + B_m \mathbf{u}$$

$$= A_p \mathbf{u}_p$$

This can then be substituted into the minimization funtion:

$$\sum_{t=0}^{N-1} \left(\|z_t\|_2^2 + \|u_t\|_2^2 \right) + \|z_N\|_2^2 = \sum_{t=0}^{N-1} \left(\left\| (A_{p,t} \mathbf{u}_{p,t})^2 \right\|_2 \right)^2 + \|u_t\|_2^2 \right) + \sum_{t=0}^{N-1} \frac{1}{N} \|z_N\|_2^2$$

$$= \sum_{t=0}^{N-1} \left(\left(\sqrt{\sum_{i=1}^{n} (A_{p,t} \mathbf{u}_{p,t})} \right)^2 + \left(\sqrt{\sum_{i=1}^{n} u_i^2} \right)^2 + \frac{1}{N} \left(\sqrt{\sum_{i=1}^{n} z_{N,i}^2} \right)^2 \right)$$

$$= \sum_{t=0}^{N-1} \left(\sum_{i=1}^{n} (A_{p,t} \mathbf{u}_{p,t})^2 + \sum_{i=1}^{n} u_i^2 + \frac{1}{N} \sum_{i=1}^{n} z_{N,i}^2 \right)$$

$$= \sum_{t=0}^{N-1} \sum_{i=1}^{n} \left((A_{p,t} \mathbf{u}_{p,t})^2 + u_i^2 + \frac{1}{N} z_{N,i}^2 \right)$$

$$= \sum_{t=0}^{N-1} \sum_{i=1}^{n} \left((A_{p,t} \mathbf{u}_{p,t})^2 + u_i^2 + \frac{1}{N} z_{N,i}^2 \right)$$

From the definition of the 2-norm for vectors and the F- norm for matrices, the following can be found:

$$= \left\| A_p \mathbf{u}_p + \mathbf{u} + \frac{1}{N} z_N \right\|_F^2$$

This can then be expanded into the follwoing matrix equivelent:

$$= \left\| A_m z_0 + B_m \mathbf{u} + \mathbf{u} + \frac{1}{N} z_N \right\|_F^2$$

$$= \left\| A_m z_0 + (B_m + I) \mathbf{u} + \frac{1}{N} z_N \right\|_F^2$$

Returning to the matrix definitions

$$A_m z_0 + (B_m + I) \mathbf{u} + \frac{1}{N} z_N = \begin{bmatrix} A & | & B & 0 & \dots & 0 \\ A^2 & | & AB & B & \dots & 0 \\ \vdots & | & \vdots & \ddots & \ddots & \vdots \\ A^N & | & A^{n-1}B & \dots & AB & B \end{bmatrix} \begin{bmatrix} z_0 \\ -- \\ \mathbf{u} \end{bmatrix} + \mathbf{u} + \frac{1}{N} z_N$$

$$= A_m z_0 + B_m \mathbf{u} + \mathbf{u} + \frac{1}{N} z_N$$

$$= A_m z_0 + (B_m + I) \mathbf{u} + \frac{1}{N} z_N$$

The minimization function can then be written as:

$$\text{minimize} \left\| A_m z_0 + (B_m + I) \mathbf{u} + \frac{1}{N} z_N I \right\|_F^2$$

which can be achieved by setting the gradient to be zero and finding the minimum solution:

$$
\begin{aligned}
0 = \nabla \left\| A_m z_0 + (B_m + I)\mathbf{u} + \frac{z_N}{N}I \right\|_F^2 \\
= \nabla \left(\sqrt{(A_m z_0)^T (A_m z_0) + ((B_m + I)\mathbf{u})^T ((B_m + I)\mathbf{u}) + \left(\frac{z_N}{N}\right)^2 I} \right)^2 \\
= \nabla \left(z_0^T A_m^T A_m z_0 + \mathbf{u}^T (B_m + I)^T (B_m + I)\mathbf{u} + \frac{z_N^2}{N^2}I \right) \\
= 2(B_m + I)\mathbf{u}
\end{aligned}
$$
$$
\mathbf{u} = 0
$$

This result states that the result that minimizes the objective function, since Z_n doesn't have an objective value, is to have $u = 0 \ \forall t \geq 0$.

Another verification can be done alternatively by defining the new vector,

$$
\mathbf{x} =
\begin{bmatrix}
z_0 \\
-- \\
z_N \\
-- \\
\mathbf{u}
\end{bmatrix},
$$

and similarly minimizing:

$$
A_x \mathbf{x} =
\begin{bmatrix}
A & | & \frac{1}{N} & | & B+1 & 0 & \dots & 0 \\
A^2 & | & \frac{1}{N} & | & AB & B+1 & \dots & 0 \\
\vdots & | & \vdots & | & \vdots & \ddots & \ddots & \vdots \\
A^N & | & \frac{1}{N} & | & A^{n-1}B & \dots & AB & B+1
\end{bmatrix}
\begin{bmatrix}
z_0 \\
-- \\
z_N \\
-- \\
\mathbf{u}
\end{bmatrix}
$$

The minimization function can then be written as:

$$
\text{minimize } \|A_x \mathbf{x}\|_F^2
$$

which can be achieved by setting the gradient to be zero and finding the minimum solution:

$$
\begin{aligned}
\nabla \|A_x \mathbf{x}\|_F^2 = 0 = \nabla \left(\sqrt{\mathbf{x}^T A_x^T A_x \mathbf{x}} \right)^2 \\
= \nabla \ \mathbf{x}^T A_x^T A_x \mathbf{x} \\
= 2\mathbf{x} A_x \\
= 2\mathbf{x} A_x \left(\frac{1}{2}\right) A_x^{-1}
\end{aligned}
$$
$$
\mathbf{x} = 0
$$

The minimum value for this case is actually found to be the trivial case of all variables set to zero, including z_0 and z_N if they were variables. This would indicate more restrictions are needed for a non-trivial solution.

References

[1] Mohamed Ali Belabbas and Xudong Chen. Optimal sensor design for secure cyber-physical systems. *IFAC-PapersOnLine*, 52(20):387 – 390, 2019. 8th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2019.