ROB 501 - Mathematics for Robotics
HW #9

Due 3 PM on Wed, Nov. 16, 2022
To be submitted on Canvas

Remark: Problems 1 and 2 develop the theory behind the "magic" formula for under determined equations. Problem 1 is hard, but the hints essentially solve it for you. When working Problem 2, it is important to realize that the Lemmas stated in Problem 1 give you everything except the Normal Equations. Problem 3 tries to unify a few concepts. If you have been hearing about "quadratic programs" (QPs), well, Problem 5 is a simple case of a QP. You also solved one in HW 1 (go back and look at Problem 6 of HW 1). We will do more advanced QPs at the end of the term.

1. Let $(\mathcal{X}, \mathbb{R}, < \cdot, \cdot >)$ be a finite-dimensional inner product space. Let $\{y_1, \cdots, y_p\}$ be a linearly independent set in $\mathcal{X}$ and let $c_1, \cdots, c_p$ be real constants. Define

$$V = \{x \in \mathcal{X} \mid \ < x, y_i >= c_i, \ 1 \leq i \leq p\}.$$

Prove the following:

(a) **Lemma 1:** There exists a unique $x_0 \in \text{span}\{y_1, \cdots, y_p\}$ such that $< x_0, y_i >= c_i, \ 1 \leq i \leq p$.

Remark: Another way of stating your result in (a) is that there exists a unique $x_0 \in \mathcal{X}$ such that

$$V \cap \text{span}\{y_1, \cdots, y_p\} = \{x_0\}.$$

(b) **Lemma 2:** Let $M = (\text{span}\{y_1, \cdots, y_p\})^{\perp}$. Then $V = x_0 + M$; in other words, $x \in V$ if, and only if, $(x - x_0) \perp \text{span}\{y_1, \cdots, y_p\}$.

(c) **Lemma 3:** There exists a unique $v^* \in V$ having minimum norm, and v^* is characterized by $v^* \perp M$ (just for emphasis, we note that the result does not say that $v^* \perp V$).

Remark: We note that $v^* \perp M \Leftrightarrow v^* \in \text{span}\{y_1, \cdots, y_p\}$ because $\mathcal{X} = M \oplus M^{\perp}$ implies that

$$M^{\perp} := \left(\text{span}\{y_1, \cdots, y_p\}^{\perp}\right)^{\perp} = \text{span}\{y_1, \cdots, y_p\}$$

Remark: We are using the standard induced norm, $||x|| =< x, x >^{1/2}$.

Remark: There exists v^* having minimum norm means $||v^*|| = \inf_{v \in V} ||v||$, and thus

$$v^* = \arg \min_{v \in V} ||v||.$$

2. Using Lemmas 1 through 3, prove the following: **Theorem:** Let $(\mathcal{X}, \mathbb{R}, < \cdot, \cdot >)$ be a finite-dimensional inner product space. Let $\{y_1, \cdots, y_p\}$ be a linearly independent set in $\mathcal{X}$ and let $c_1, \cdots, c_p$ be real constants. Define $V = \{x \in \mathcal{X} \mid \ < x, y_i >= c_i,\ 1 \leq i \leq p\}$. Then there exists a unique $v^* \in V$ such that

$$v^* = \arg \min_{v \in V} ||v||.$$

Moreover, $v^* = \sum_{i=1}^{p} \beta_i y_i$, where the β_i's satisfy the normal equations

$$\begin{bmatrix} < y_1, y_1 > & < y_2, y_1 > & \cdots & < y_p, y_1 > \\ < y_1, y_2 > & < y_2, y_2 > & \cdots & < y_p, y_2 > \\ \vdots & \vdots & \ddots & \vdots \\ < y_1, y_p > & < y_2, y_p > & \cdots & < y_p, y_p > \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}$$

Remark: What you have just solved is a special case of a Quadratic Program, typically called a **QP** for short.

3. This problems seeks to relate several concepts we have seen in the course. Suppose that X and Y are jointly distributed normal random variables with

$$\mu =: \begin{bmatrix} \mathcal{E}\{X\} \\ \mathcal{E}\{Y\} \end{bmatrix} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} \quad \text{and} \quad \Sigma = \text{cov}(\begin{bmatrix} X \\ Y \end{bmatrix}, \begin{bmatrix} X \\ Y \end{bmatrix}) = \begin{bmatrix} P & PC^\top \\ CP & CPC^\top + Q \end{bmatrix}$$

   (a) Compute the mean and covariance of X conditioned on $Y = y$, using Fact 1 in the handout on Gaussian Random Vectors, and compare to the formula for the MVE given in HW8 Problem 6.

   (b) Compute the Schur complement of $CPC^\top + Q$ in Σ and compare to the covariance of X conditioned on Y.

   Remark: Why is this interesting? The Minimum Variance Estimator (MVE) was derived using the Projection Theorem. We computed $\widehat{x}$ as the orthogonal projection of x onto the measurement, y. From this problem, we get the hint that when working with Gaussian random vectors, conditional expectations are orthogonal projections. If you take EECS 564 (Estimation and Detection), this fact is actually proven! I hope this helps to bring together the various estimation schemes.

4. Consider the finite dimensional vector space $(\mathcal{X}, \mathbb{R})$, where $\mathcal{X} = \text{span}\{1, t, t^2, \sin(\pi t)\}$. Equip it with the inner product $< f, g >:= \int_0^2 f(\tau) g(\tau) d\tau$. When working the problem, feel free to use MATLAB to compute any required integrals, and you may compute them symbolically or numerically, as you wish.

   (a) Find the vector of minimum norm that satisfies $< f, t >= 2$.

   (b) Find the vector of minimum norm that satisfies $< f, t >= 2$ and $< f, \sin(\pi t) >= \pi$.

5. **Underdetermined Equations:** We consider $Ax = b$, where $b \in \mathbb{R}^p$, $x \in \mathbb{R}^n$, $n > p$. Use the normal equations derived in Problem 2 to solve the following problems. The key is to interpret the rows of $Ax = b$ in terms of inner product conditions that look like $< x, y_i >= c_i$.

   (a) We assume an inner product on $\mathbb{R}^n$ defined by $< x, z >:= x^\top z$, and thus $||x|| = (x^\top x)^{1/2}$. Show that if the rows of A are linearly independent, then

$$\widehat{x} := \arg \min_{Ax = b} ||x||$$

   is given by $\widehat{x} = A^\top \left(AA^\top\right)^{-1} b$

(b) We assume an inner product on $\mathbb{R}^n$ defined by $< x, z >:= x^\top Q z$, where $Q \succ 0$, and thus $||x|| = (x^\top Q x)^{1/2}$. Show that if the rows of A are linearly independent, then

$$\widehat{x} := \arg \min_{Ax=b} ||x||$$

is given by $\widehat{x} = Q^{-1} A^\top \left(A Q^{-1} A^\top \right)^{-1} b$.

Remark: A QP is an optimization problem of the form

$$\widehat{x} := \quad \begin{array}{c} \arg \min \\ A_{eq}x = b_{eq} \\ A_{in}x \leq b_{in} \end{array} \quad ||x||^2.$$

The key addition is that inequality constraints, $A_{in}x \leq b_{in}$, can also be included. We'll talk more about this the last day of lecture.

6. We did one version of the QR factorization in lecture (we assumed A had linearly independent columns). Scan MATLAB's documentation of the `qr` function and note its *economy* version, `[Q,R]=qr(A,0)`. Using Gram Schmidt, compute the QR factorization of the matrix A below "by hand" (you can do any vector multiplications and such in MATLAB[1]). Compare your answer to MATLAB's `[Q,R]=qr(A)` and `[Q,R]=qr(A,0)` and discuss similarities and differences.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

[1] This is just to encourage you to do one QR factorization yourself without relying on MATLAB's built-in function.

Hints

Hints: Prob. 1 The set V is similar to looking at the set of solutions of a linear equation $\{x \in \mathbb{R}^n \mid Ax = b\}$. When $b = 0$, you have a subspace. When $b \neq 0$, you have the same subspace translated by x_0 where $Ax_0 = b$. If this confuses you, then ignore it. I am just trying to give you some intuition! The real hints follow:

(a) Express $x_0 = \sum_{i=1}^{p} \alpha_i y_i$, a linear combination of the y_i's, and then plug into the conditions for $x_0 \in V$. You will get something that looks exactly like the normal equations! You will see a Gram matrix. You will recall that a Gram matrix is invertible if, and only if, the set $\{y_1, \cdots, y_p\}$ is linearly independent.

(b) We have that $x_0 \in V$. Suppose that $x \in V$. Then $< x, y_i >= c_i =< x_0, y_i >$ for $1 \leq i \leq p$. Now ask yourself about $< x - x_0, y_i >$ and see what that tells you about the relation of $x - x_0$ to $\text{span}\{y_1, \cdots, y_p\}$.

(c) This one is more conceptual than the previous two parts. We now know what elements in V look like. In particular, $v \in V$ if, and only if, $v = x_0 + m$ for $m \in M$. Because M is a subspace, we can also write this as $v \in V$ if, and only if, $v = x_0 - m$ for $m \in M$ (because $m \in M \Leftrightarrow -m \in M$). Hence, we have

$$\inf_{v \in V} ||v|| = \inf_{m \in M} ||x_0 + m|| = \inf_{m \in M} ||x_0 - m|| = d(x_0, M).$$

From the Projection Theorem, you know a lot about the right side of the above string of equalities. Use this knowledge to characterize the optimal $m^* \in M$, and then apply that knowledge to $v^* = x_0 - m^*$.

Hints: Prob. 2 Almost everything has been proved in Prob. 1. Here, it is mainly a matter of assembling the pieces into the final "beautiful" form of the answer. Note that in Prob. 1, part(a), you derived equations that look remarkably like those in the Theorem of Prob. 2. However, when you were working (a), you had no idea that the particular solution you were computing was in fact the element in V of minimum norm. Part (c) of Prob. 1 is what allows you to make that connection. Because we are working with $\mathcal{F} = \mathbb{R}$, $< y_i, y_j >=< y_j, y_i >$, and thus the Gram matrix is symmetric. Hence, if you end up with the ij component as $< y_i, y_j >$ instead of $< y_j, y_i >$, you are not wrong!

Hints: Prob. 4 Apply result from Problem 2. You may enjoy the symbolic toolbox for doing the calculations.

```
syms t pi
f=t^2;
g=exp(pi*t);
a=0; b=2;
G11=int(f*f,a,b);
G11=simple(G11)
```

Commands such as `inv` and `numden` work in the symbolic toolbox. Also checkout `simplify` and `pretty`.

Hints: Prob. 5 Decompose A by its rows,

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{bmatrix}.$$

Define $\tilde{A} = AQ^{-1}$ (why is Q invertible?), and define $v_i \in \mathbb{R}^n$ by

$$v_i = (a_i Q^{-1})^\top = Q^{-1} a_i^\top,$$

where we have used the fact that Q is symmetric. Show that

$$Ax = b \Leftrightarrow < v_i, x >= b_i, \quad 1 \leq i \leq p.$$

Now, work out those normal equations from Problem 2! If the above derivation confuses you, set $Q = I$ and you will get it! That is why I broke the problem into two parts.

Remark: Squaring the norm does not change the answer[2] to the optimization problem because it is a strictly monotonically increasing function.

$$\widehat{x} := \arg \min_{Ax=b} ||x|| = \arg \min_{Ax=b} ||x||^2 = \arg \min_{Ax=b} x^\top Q x.$$

The corresponding QP is always written using the form on the right.

Hints: Prob. 6 Nothing exciting here. Just pointing out that there exist multiple versions of the QR factorization.

[2] Yes, the value is squared, but the argument achieving the minimum, i.e., the $\widehat{x}$, does not change.