ELEC 3035: Practice problems for part 1

Lecturer: Ivan Markovsky

1. *Angle between vectors and length of a vector* Verify that the triangle in $\mathbb{R}^2$ with vertexes $(1/2, 1/2)$, $(2, -1)$, and $(4, 4)$ is a right triangle and verify that the Pythagorean theorem holds for it.

2. *Inverse of a 2×2 matrix and related graph* Find the inverse of the matrix $A = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]$. Visualize the mapping $u \overset{A}{\mapsto} y := Au \overset{A^{-1}}{\mapsto} z := A^{-1}y$ by drawing a graph from u_1, u_2 to y_1, y_2 and from y_1, y_2 to z_1, z_2. Since $z = u$, the graph from u_1, u_2 to z_1, z_2 is simple: $\begin{matrix} u_1 \mapsto u_1 \\ u_2 \mapsto u_2 \end{matrix}$. Think of $y := Au$ as a *decoder* and of $u = A^{-1}y$ as an *encoder*. (For larger matrices you have to relay on a computer for doing the coding and decoding operations for you.)

3. *Distance to a subspace* The 2-norm distance from a point $d \in \mathbb{R}^n$ to a set $\mathscr{B} \subset \mathbb{R}^n$ is defined as

$$\operatorname{dist}(d, \mathscr{B}) := \min_{\widehat{d} \in \mathscr{B}} \|d - \widehat{d}\|_2, \tag{1}$$

i.e., $\operatorname{dist}(d, \mathscr{B})$ is the shortest distance from d to a point $\widehat{d}$ in $\mathscr{B}$. A vector $\widehat{d}^*$ that achieves the minimum of (1) (it need not be unique) is a point in $\mathscr{B}$ that is closest to d.

Next we consider the special case when $\mathscr{B}$ is a subspace.

(a) Let $\mathscr{B} = \operatorname{image}(a) = \{\, \alpha a \mid \alpha \in \mathbb{R} \,\}$. Explain how to find $\operatorname{dist}\big(d, \operatorname{image}(a)\big)$. Find

$$\operatorname{dist}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \operatorname{image}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \right)$$

and sketch the solution. Note that the best approximation $\widehat{d}^*$ of d in $\operatorname{image}(a)$ is the orthogonal projection of y onto $\operatorname{image}(a)$.

(b) Let $\mathscr{B} = \operatorname{image}(P)$. Explain how to find $\operatorname{dist}(d, \mathscr{S})$. (You can assume that P is full column rank. Argue that this can be done without loss of generality.)

(c) Let $\mathscr{B} = \ker(R)$. Explain how to find $\operatorname{dist}(d, \mathscr{S})$. (You can assume that R is full row rank. Argue that this can be done without loss of generality.)

(d) In the case when $\mathscr{B}$ is a subspace, is a solution $\widehat{d}^*$ of (1) always unique?

(e) Prove that when $\mathscr{B}$ is a subspace, $y - \widehat{y}^*$ is orthogonal to $\mathscr{S}$. Is the converse true, *i.e.*, is it true that if for some $\widehat{y}$, $y - \widehat{y}$ is orthogonal to $\mathscr{S}$, then $\widehat{y} = \widehat{y}^*$?

Solution:

(b) Using the image representation $\operatorname{image}(P)$ of the subspace $\mathscr{B}$, the distance computation problem (1) is equivalent to the standard least squares problem

$$\operatorname{dist}(d, \mathscr{B}) := \min \|d - \widehat{d}\|_2, \quad \text{subject to} \quad \widehat{d} = Pl.$$

Therefore, assuming that P is full column rank, the best approximation is

$$\widehat{d}^* = P(P^\top P)^{-1} P^\top d \tag{2}$$

and (after some work) the distance of d to $\mathscr{B}$ is

$$\operatorname{dist}(d, \mathscr{B}) = \|d - d^*\|_2 = \sqrt{d^\top \big(I - P(P^\top P)^{-1} P^\top\big) d}. \tag{3}$$

The assumption that "P is full column rank" can be done without loss of generality because there are aways full column rank P's such that $\operatorname{image}(P) = \mathscr{B}$ (choose any basis for $\mathscr{B}$).

(c) Using the kernel representation $\ker(R)$ of the subspace $\mathscr{B}$, the distance computation problem (1) is equivalent to the problem

$$\operatorname{dist}(d,\mathscr{B}) := \min \|d - \widehat{d}\|_2, \quad \text{subject to} \quad \widehat{R}d = 0.$$

As written, this problem is not a standard problem, however, with the change of variables $\widetilde{d} := d - \widehat{d}$ it can be rewritten as an equivalent ordinary least norm problem

$$\operatorname{dist}(d,\mathscr{B}) := \min \|\widetilde{d}\|_2, \quad \text{subject to} \quad R\widetilde{d} = Rd.$$

Therefore, assuming that R is full row rank,

$$\widetilde{d}^* = R^\top (RR^\top)^{-1} Rd$$

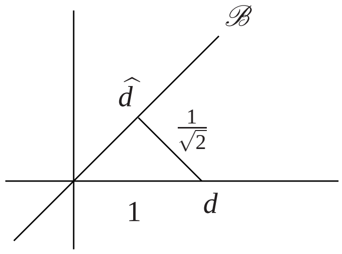
and

$$\operatorname{dist}(d,\mathscr{B}) = \|\widetilde{d}^*\|_2 = \sqrt{d^\top R^\top (RR^\top)^{-1} Rd}. \tag{4}$$

Again, the assumption that R is full row rank is without loss of generality because there are full row rank matrices R, such that $\ker(R) = \mathscr{B}$.

(a) Substituting $d = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$ and $P = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ in (3), we have

$$\operatorname{dist}\left(\begin{bmatrix}1\\0\end{bmatrix}, \operatorname{image}\left(\begin{bmatrix}1\\1\end{bmatrix}\right)\right) = \sqrt{\begin{bmatrix}1 & 0\end{bmatrix}\left(\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}1\\1\end{bmatrix}\left(\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right)^{-1}\begin{bmatrix}1 & 1\end{bmatrix}\right)\begin{bmatrix}1\\0\end{bmatrix}}$$
$$= \sqrt{\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1/2 & -1/2\\-1/2 & 1/2\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}} = 1/\sqrt{2}$$



(d) Yes, as shown in part 3b, $\widehat{d}^*$ is unique (and can be computed by (2)).

(e) The vector $d - \widehat{d}^*$ is orthogonal to the subspace $\mathscr{B}$ if and only if $d - \widehat{d}^*$ is orthogonal to all vectors of a basis of $\mathscr{B}$. Using (2) and the basis P, we have

$$(\widehat{d} - d^*)^\top P = d^\top \big(I - P(P^\top P)P^\top\big)P = 0,$$

which shows that is $d - \widehat{d}^*$ is orthogonal to $\mathscr{B}$.

The converse statement ($d - \widehat{d}$ being orthogonal to $\mathscr{B}$ implies that $\widehat{d}$ is the closest point in $\mathscr{B}$ to d) is true but is more difficult to show. It completes the proof of what is known as the "orthogonality principle"—a point $\widehat{d}$ is an optimal approximation of a point d in a subspace $\mathscr{B}$ if and only if the approximation error $d - \widehat{d}$ is orthogonal to $\mathscr{B}$.

□

4. *Distance to an affine space* Consider again the distance $\operatorname{dist}(d,\mathscr{B})$ defined in (1). In this problem, we consider the case when $\mathscr{B}$ is an affine set, *i.e.*, $\mathscr{B} = \mathscr{S} + a$, where $\mathscr{S}$ is a subspace and a is a shift.

(a) Explain how to reduce the problem of computing the distance from a point to an affine space to an equivalent problem of computing the distance to a subspace.

(b) Find

$$\operatorname{dist}\left(\begin{bmatrix}0\\0\end{bmatrix}, \ker(\begin{bmatrix}1 & 1\end{bmatrix}) + \begin{bmatrix}1\\2\end{bmatrix}\right)$$

and sketch the solution.

Solution:

- The problem of computing $\operatorname{dist}(d, \mathscr{A})$ reduces to an equivalent problem of computing the distance of a point to a subspace by the change of variables $d' := d - a$

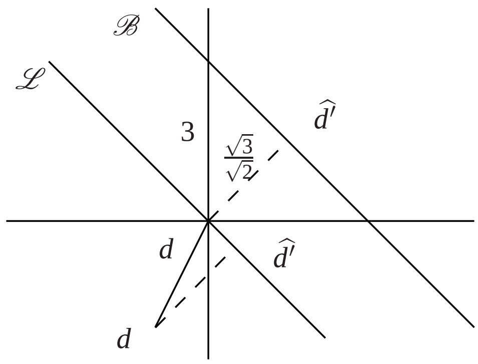
$$\operatorname{dist}(d, \mathscr{B}) = \min_{\widehat{d} \in \mathscr{B}} \|d - \widehat{d}\|_2 = \min_{\widehat{d'} \in \mathscr{L}} \|d' - \widehat{d'}\|_2 = \operatorname{dist}(d', \mathscr{L}).$$

- Using the change of variables argument we have

$$\operatorname{dist}\left(\begin{bmatrix}0\\0\end{bmatrix}, \ker(\begin{bmatrix}1 & 1\end{bmatrix}) + \begin{bmatrix}1\\2\end{bmatrix}\right) = \operatorname{dist}\left(-\begin{bmatrix}1\\2\end{bmatrix}, \ker(\begin{bmatrix}1 & 1\end{bmatrix})\right).$$

  Then using (4) we have

$$\operatorname{dist}\left(-\begin{bmatrix}1\\2\end{bmatrix}, \ker(\begin{bmatrix}1 & 1\end{bmatrix})\right) = \sqrt{\begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\left(\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\right)^{-1}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}} = \sqrt{9/2}.$$



□

5. *Hand computation of eigenvalues and eigenvectors of a* 2×2 *matrix* Find the eigenvalues and a set of linearly independent eigenvectors of the matrix $\left[\begin{smallmatrix}0 & 1\\1 & 1\end{smallmatrix}\right]$.

   Answer:

$$\begin{array}{lll} \text{Eigenvalues:} & \lambda_1 = \frac{1+\sqrt{5}}{2} & \lambda_2 = \frac{1-\sqrt{5}}{2} \\ \text{Eigenvectors:} & v_1 = \alpha \begin{bmatrix}1\\\lambda_1\end{bmatrix}, \alpha_1 \in \mathbb{R} & v_2 = \alpha_2 \begin{bmatrix}1\\\lambda_2\end{bmatrix}, \alpha_2 \in \mathbb{R} \end{array}$$

6. *Fibonacci numbers* The Fibonacci numbers $f(0), f(1), f(2), \ldots$ are defined by

$$f(0) = 1, \quad f(1) = 1, \quad \text{and} \quad f(t+1) = f(t) + f(t-1), \quad \text{for } t = 2, 3, \ldots. \tag{FN}$$

   Find the 50th Fibonacci numbers $f(50)$.

   Solution:

   A polynomial representation of (FN) is $r(\sigma)f = 0$, where $r(z) = z^2 - z - 1$. The general solution is of the form

$$f(t) = c_1 z_1^t + c_2 z_2^t$$

where $z_1 = \frac{1+\sqrt{5}}{2}$ and $z_2 = \frac{1-\sqrt{5}}{2}$ are the roots of $r(z)$ (the poles of the system) and c_1 and c_2 are constants depending on the initial conditions. In order to find c_1 and c_2, we solve the system

$$\begin{array}{l} f(0) = c_1 z_1^0 + c_2 z_2^0 \\ f(1) = c_1 z_1^1 + c_2 z_2^1 \end{array} \qquad \Longleftrightarrow \qquad \begin{bmatrix} 1 & 1 \\ z_1 & z_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \Longleftrightarrow \qquad \begin{array}{l} c_1 = \frac{z_2 - 1}{z_2 - z_1} \\ c_2 = \frac{1 - z_1}{z_2 - z_1} \end{array}$$

Finally, we have

$$f(50) = \frac{z_2 - 1}{z_2 - z_1} z_1^{50} + \frac{1 - z_1}{z_2 - z_1} z_2^{50},$$

which is an explicit expression (since z_1 and z_2 are known numbers) and can be evaluated to find that

$$f(50) = 20365011074.$$

□

7. *Harmonic oscillator* The differential equation defining the behaviour of a harmonic oscillator is

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} y = -ky,$$

where k is a positive constant. (A physical example of a harmonic oscillator is a unit mass attached to a spring, in which case k is the spring constant. You have seen harmonic oscillators also in circuit theory.) Find a state space representation of the harmonic oscillator. Give a formula for the trajectories of the system, starting from a given initial condition.

(See, Lecture 3, page 16.)

8. *Feedforward control* Given a static system $\mathscr{B}_2 = \{ (u,y) \mid y = A_2 u \}$ (the plant), explain how to find a system $\mathscr{B}_1 = \{ (r,u) \mid u = A_1 r \}$ (feedforward controller), such that the series connection of $\mathscr{B}_1$ and $\mathscr{B}_2$ (the controlled system) matches or is as close as possible to a given system $\mathscr{B}_3 = \{ (r,y) \mid y = A_3 r \}$ (reference model). You can assume that the matrix A_2 is full rank.

Solution: The controlled system is

$$\begin{aligned} \mathscr{B}_\mathrm{c} &= \{ (r,y) \mid \text{ there is } u \text{ such that } (r,u) \in \mathscr{B}_1 \text{ and } (u,y) \in \mathscr{B}_2 \} \\ &= \{ (r,y) \mid \text{ there is } u \text{ such that } u = A_1 r \text{ and } y = A_2 u \} \\ &= \{ (r,y) \mid y = \underbrace{A_2 A_1}_{A_\mathrm{c}} r \} = \{ (r,y) \mid y = A_\mathrm{c} r \}. \end{aligned}$$

Then the problem of matching the reference model with the controlled system is equivalent to the problem of solving the linear system of equations

$$A = A_2 A_1$$

for the unknown A_1. Since A_2 is full rank, solution A_1 exists, for *any* given A, if and only if the system of equations has as many unknowns as equations or is overdetermined. Let $m = \dim(u)$, $p = \dim(y)$, and $d = \dim(r)$. The number of equations is pd and the number of unknowns is md, so we distinguish the following cases:

- $m = p$ — there is a unique solution $A_1 = A_2^{-1} A$,
- $m > p$ — there are infinitely many solutions, a particular solution is the least norm one, $A_{1,\mathrm{ln}} = A_2^\top (A_2 A_2^\top)^{-1} A$,
- $p > m$ — there is no solution, a least squares approximate solution is, $A_{1,\mathrm{ls}} = (A_2^\top A_2)^{-1} A_2^\top A$.

□

9. *Questions on controllability*

   (a) Give an example of uncontrollable first order system (we will call this system $\mathscr{B}_1$) and an example of uncontrollable second order system (we will call this system $\mathscr{B}_2$).

(b) Sketch the controllable subspaces of $\mathscr{B}_1$ and $\mathscr{B}_2$.

(c) For $\mathscr{B}_2$, choose a specific state $x_{\text{des}} \in \mathbb{R}^2$ that is not reachable from the origin (in any number of time steps) and find the closest state $\widehat{x}_{\text{des}}$ to x_{des} that is reachable from the origin.

(d) For $\mathscr{B}_2$, derive a control input that transfers $x(0) = 0$ to $x(2) = \widehat{x}_{\text{des}}$.

10. *Questions on observability* Consider a given autonomous linear system $\mathscr{B}$ of order n with p outputs and a given trajectory $y_{\text{d}} = \big(y_{\text{d}}(1), \dots, y_{\text{d}}(T)\big)$ of $\mathscr{B}$. You can assume that $\mathscr{B}$ is specified by an observable state space representation.

    (a) Is it always possible to predict the future of y_{d} from $T \geq n$ samples?

    (b) Are there cases when the future of y_{d} can be predicted from $T < n$ samples?

    (c) In the cases when the future of y_{d} can not be predicted from the given data, would it help to use future data points $y_{\text{d}}(t_1, \dots, t_K)$, $n < t_1 < \cdots < t_K$. If so, how many? Do they need to be sequential, *i.e.*, $t_{i+1} = t_i + 1$?

11. *[Lue79, Chapter 2, Problem 2]* A bank offers 7% annual interest. What would be the overall annual rate if the 7% interest were compounded quarterly?

    Solution: Let $y(k)$ denote the amount in the account at the beginning of season k and the bank pays interest at the end of each season. If the 7% interest were compounded quarterly, then the quarter interest is 7/4% and the account balance is governed by

    $$y(k+1) = (1 + 7/4\%)y(k).$$

    For an year, suppose $y(k)$ is the amount in the account at the beginning of this year, then $y(k+4)$ is the amount in the account at the beginning of following year. In order to know the overall annual rate, we should specify the relationship between $y(k)$ and $y(k+4)$. We have

    $$\begin{aligned} y(k+4) &= (1+7/4\%)y(k+3) = (1+7/4\%)^2 y(k+2) \\ &= (1+7/4\%)^3 y(k+1) = (1+7/4\%)^4 y(k). \end{aligned}$$

    Thus, the annual rate is

    $$(1+7/4\%)^4 - 1 = 7.1859\%.$$

    □

12. *[Lue79, Chapter 2, Problem 5]* Find the second order linear homogeneous difference equation which generates the sequence 1, 2, 5, 12, 29, 70, 169. What is the limiting ratio of consecutive terms?

    Solution: By observing the sequence we find the relationship among three consecutive terms is

    $$\begin{aligned} 5 &= 2 \cdot 2 + 1 \\ 12 &= 2 \cdot 5 + 2 \\ 29 &= 2 \cdot 12 + 5 \\ 70 &= 2 \cdot 29 + 12 \\ 169 &= 2 \cdot 70 + 29 \end{aligned}$$

    This relation can be written as a second-order linear homogeneous difference equation

    $$y(k+2) = 2y(k+1) + y(k).$$

    By dividing both sides of the equation by $y(k+1)$, the equation becomes

    $$\frac{y(k+2)}{y(k+1)} = 2 + \frac{y(k)}{y(k+1)}.$$

Here $\frac{y(k+1)}{y(k)}$ is the ratio of two consecutive term. When $k \to \infty$, the ratio converges to a constant which defines as a. Therefore

$$\frac{y(k+2)}{y(k+1)} = \frac{y(k+1)}{y(k)} = a, \qquad k \to \infty.$$

The limiting ratio a satisfy an equation

$$a = 2 + \frac{1}{a} \quad \Longrightarrow \quad a^2 = 2a + 1 \quad \Longrightarrow \quad a = 1 \pm \sqrt{2}.$$

However, all terms of the sequence are positive so that the ratio of consecutive terms is positive. Therefore, the limiting ratio $a = 1 + \sqrt{2}$. □

13. *[Lue79, Chapter 2, Problem 10]* Consider the second order difference equation

$$y(k+2) - 2ay(k+1) + a^2 y(k) = 0.$$

Its characteristic polynomial has both roots equal to $\lambda = a$.

(a) Show that both

$$y(k) = a^k \qquad \text{and} \qquad y(k) = ka^k$$

are solutions.

(b) Find the solutions of this equation that satisfies the auxiliary conditions $y(0) = 1$ and $y(1) = 0$.

Solution:

(a) To check $y(k) = a^k$, we note that $y(k+2) = a^{k+2}$ and $y(k+1) = a^{k+1}$

$$a^{k+2} - 2a \cdot a^{k+1} + a^2 \cdot a^k = a^{k+2} - 2a^{k+2} + a^{k+2} = 0.$$

Thus, $y(k) = a^k$ is a solution. Checking $y(k) = ka^k$ we have

$$\begin{aligned}(k+2)a^{k+2} - 2a(k+1)a^{k+1} + a^2 ka^k &= (k+2)a^{k+2} - 2(k+1)a^{k+2} + ka^{k+2} \\ &= (k+2-2k-2+k)a^{k+2} = 0.\end{aligned}$$

Thus, $y(k) = ka^k$ is also a solution.

(b) A second-order linear difference equation has two degrees of freedom in its general solution, *i.e.*, two linearly independent solutions can form a fundamental set of solutions. We have found two solutions a^k and ka^k. It is easy to prove that these two solutions are linear independent. Because we can't find two constant c_1 and c_2 at least one of which is nonzero to satisfy

$$c_1 a^k + c_2 ka^k = 0$$

for all $k = 0, 1, 2, \ldots, N$. Thus, any solution $y(k)$ can be expressed as a linear combination of the fundamental set of solutions a^k and ka^k.

$$y(k) = c_1 a^k + c_2 ka^k,$$

where c_1 and c_2 are constant. Using the conditions $y(0) = 1$ and $y(1) = 0$, we find c_1 and c_2

$$\begin{aligned} y(0) = c_1 a^0 = 1 \quad &\Longrightarrow \quad c_1 = 0 \\ y(1) = c_1 a + c_2 a = a + c_2 a = 0 \quad &\Longrightarrow \quad c_2 = -1. \end{aligned}$$

Thus, the solution to this equation is

$$y(k) = a^k - ka^k.$$

□

References

[Lue79] D. G. Luenberger. *Introduction to Dynamical Systems: Theory, Models and Applications*. John Wiley, 1979.