ELEC 3035: Tutorial on linear algebra **Lecturer: Ivan Markovsky**

1. *Identification of a linear function* Find conditions under which a linear function $f : \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{p}}$ can be identified from given input/output pairs $(u^{(1)}, y^{(1)}), \dots, (u^{(N)}, y^{(N)})$, such that $y^{(i)} = f(u^{(i)})$, for $i = 1, \dots, N$. Assuming that the identifiability conditions are satisfied, explain how f can be found algorithmically in terms of standard linear algebra operations. Check the identifiability conditions on the data

   (a) $u^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; y^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad u^{(2)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \; y^{(2)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix};$

   (b) $u^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; y^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad u^{(2)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \; y^{(2)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix};$

   and if they are satisfied, find the corresponding functions f. (Note that these are two different examples with possibly different functions f.)

   Solution: Since f is linear, there is a matrix $A \in \mathbb{R}^{\mathtt{p} \times \mathtt{m}}$, such that $y = f(u) = Au$.

   $$y^{(i)} = f(u^{(i)}), \quad \text{for } i = 1, \dots, N \qquad \iff \qquad \underbrace{\begin{bmatrix} y^{(1)} & \cdots & y^{(N)} \end{bmatrix}}_{Y \in \mathbb{R}^{\mathtt{p} \times N}} = A \underbrace{\begin{bmatrix} u^{(1)} & \cdots & u^{(N)} \end{bmatrix}}_{U \in \mathbb{R}^{\mathtt{m} \times N}} \qquad (*)$$

   The equation $Y = AU$ is a linear system for A. We know in advance that a solution exists. (Why?) The identifiability condition for f is the condition that the system $Y = AU$ has a unique solution (which therefore must be the "correct" one). This is the case if and only if U is full row rank.

   The last statement follows from more general results about uniqueness of solution of linear systems. To see why it is true, note that the number of linearly independent rows is equal to the number of linearly independent columns. U being full row rank implies that there are $\mathtt{m}$ linearly independent rows, so that there must be $\mathtt{m}$ linearly independent columns. Let U' be the matrix formed by these columns and let Y' be the matrix formed by the corresponding columns of Y. We have $Y' = AU'$ with $U' \in \mathbb{R}^{\mathtt{m} \times \mathtt{m}}$ nonsingular. Therefore, the solution is unique and is given by $A = Y'U'^{-1}$.

   (a) In this example $U = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $Y = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. $\det(U) = -1 - 1 = -2 \neq 0$, so that U is nonsingular. Therefore,

   $$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

   (For the last calculation, the following formula is helpful $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Check it.) The function $f(u) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} u$ flips the the elements of u.

   (b) In this example, $U = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ is not full row rank, so that the function is not identifiable from the data. To show this we display two functions $f'(u) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} u$ and $f''(u) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} u$ that satisfy the relations $y^{(i)} = f'(u^{(i)})$ $y^{(i)} = f''(u^{(i)})$, $i = 1, 2$. (Check this.) There is no way to distinguish between f' and f''. Actually, there are infinite many linear functions that satisfy the data. (Why?) Show one more.

   □

2. *Identification of an affine function* An affine function is a linear function plus a constant, i.e.,

$$f : \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{p}} \text{ is affine} \quad \iff \quad \text{there are } A \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}} \text{ and } b \in \mathbb{R}^{\mathtt{p}} \text{ such that } f(u) = Au + b.$$

Find conditions under, which an affine function $f : \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{p}}$ can be identified from given input/output pairs $(u^{(1)}, y^{(1)}), \ldots, (u^{(N)}, y^{(N)})$, such that $y^{(i)} = f(u^{(i)})$, for $i = 1, \ldots, N$. Under the identifiability conditions, explain how f can be computed algorithmically in terms of standard linear algebra operations. Check the identifiability conditions on the data

(a) $u^{(1)} = \begin{bmatrix}1\\1\end{bmatrix}, y^{(1)} = \begin{bmatrix}2\\2\end{bmatrix}, \quad u^{(2)} = \begin{bmatrix}1\\-1\end{bmatrix}, y^{(2)} = \begin{bmatrix}0\\2\end{bmatrix}, \quad u^{(3)} = \begin{bmatrix}0\\1\end{bmatrix}, y^{(3)} = \begin{bmatrix}2\\1\end{bmatrix};$

(b) $u^{(1)} = \begin{bmatrix}1\\1\end{bmatrix}, y^{(1)} = \begin{bmatrix}2\\2\end{bmatrix}, \quad u^{(2)} = \begin{bmatrix}2\\2\end{bmatrix}, y^{(2)} = \begin{bmatrix}3\\3\end{bmatrix}, \quad u^{(3)} = \begin{bmatrix}0\\1\end{bmatrix}, y^{(3)} = \begin{bmatrix}2\\1\end{bmatrix};$

and if they are satisfied, find the corresponding functions f. (Note that these are two different examples with possibly different functions f.)

Solution: This problem is reduced to the previous one by grouping together the unknowns A and b

$$y = Au + b = \underbrace{\begin{bmatrix}A & b\end{bmatrix}}_{A_{\text{ext}}} \underbrace{\begin{bmatrix}u\\1\end{bmatrix}}_{u_{\text{ext}}} = A_{\text{ext}} u_{\text{ext}},$$

which give us a linear function in terms of the extended input u_{ext}. Applying the result of Problem 1, we have that the affine function is identifiable from the data if and only if the matrix

$$U_{\text{ext}} := \begin{bmatrix} u^{(1)} & \cdots & u^{(N)} \\ 1 & \cdots & 1 \end{bmatrix}$$

is full row rank.

(a) We have

$$U_{\text{ext}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

We need to check whether this 3×3 matrix is full row rank. One way to do this is by computing the determinant and checking whether it is equal to 0. If you have forgotten the formula (which you need not remember), you can try to find the inverse. This latter approach is systematic and as a result will bring you very close to the solution of the problem, provided U_{ext} is nonsingular.

Please recall from your first year linear algebra course how to find the inverse of a matrix, using Gaussian elimination. Here is the result of applying the procedure

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & -2 & 0 & 0 & 1 & -1 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & -2 & 0 & 0 & 1 & -1 \\ 0 & 1 & 1 & -1 & -1/2 & 3/2 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & -2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1/2 & -1/2 \\ 0 & 1 & 0 & 0 & -1/2 & 1/2 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right] \quad \Longrightarrow \quad U_{\text{ext}}^{-1} = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \\ -1 & 0 & 1 \end{bmatrix}$$

Since the procedure succeeded the matrix U_{ext} is full row rank and as a byproduct of the procedure we obtained U_{ext}^{-1}.

$$A_{\text{ext}} = Y U_{\text{ext}}^{-1} = \begin{bmatrix} 2 & 0 & 2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Therefore the affine function that generated the data is

$$f(u) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} u + \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(b) In this example,

$$U_{\text{ext}} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

which turns out to be full rank as well. Therefore, following the same procedure as in 2a will give us the data generating affine function. The final answer is the same as in 2a. (I was aiming to get a rank deficient matrix as in Problem 1.)

□

3. *Linear least squares approximation of a nonlinear function* Consider a nonlinear function $f : \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{p}}$ and input/output pairs $(u^{(1)}, y^{(1)}), \ldots, (u^{(N)}, y^{(N)})$, such that $y^{(i)} = f(u^{(i)})$, for $i = 1, \ldots, N$. Find a linear function $\widehat{f} : \mathbb{R}^{\mathtt{m}} \to \mathbb{R}^{\mathtt{p}}$ that is as close as possible to the given points in the least squares sense, i.e., a function $\widehat{f}$ that minimizes the criterion $\sum_{i=1}^{N} \|y^{(i)} - \widehat{f}(u^{(i)})\|_2^2$. Such a function $\widehat{f}$ is referred to as a linear least squares approximation of a nonlinear function (at the points $(u^{(1)}, y^{(1)}), \ldots, (u^{(N)}, y^{(N)})$). Is $\widehat{f}$ unique? Find the linear least squares approximation $\widehat{f}$ of $f(u) = u^2$ at $u^{(1)} = 0$, $u^{(2)} = 1$, $u^{(3)} = 2$.

Solution: The uniqueness condition is the same as in Problem 1, i.e., $\operatorname{rank}(\begin{bmatrix} u^{(1)} & \cdots & u^{(N)} \end{bmatrix}) = \mathtt{m}$. Under this condition the linear least squares approximation follows from $(*)$ and is given by

$$\widehat{A} = YU^\top (UU^\top)^{-1}.$$

The linear least squares approximation of $f(u) = u^2$ at $u^{(1)} = 0$, $u^{(2)} = 1$, $u^{(3)} = 2$ is

$$\widehat{A} = \begin{bmatrix} 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \left(\begin{bmatrix} 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right)^{-1} = 9/5.$$

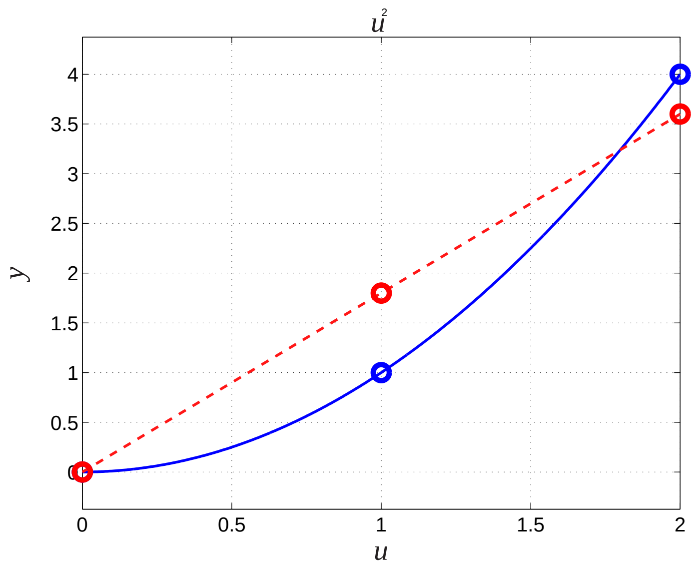

Figure 1: Least squares linear approximation $\widehat{f}(u) = 9/5u$ (red bashed line) of $f(u) = u^2$ (blue line) at the points $u^{(1)} = 0$, $u^{(2)} = 1$, $u^{(3)} = 2$.

□

4. *Polynomial interpolation* A scalar polynomial function of degree n is a function $p : \mathbb{R} \to \mathbb{R}$,

$$p(u) := p_0 u^0 + p_1 u^1 + \cdots + p_n u^n, \quad p_n \neq 0$$

where $p_0, p_1, \dots, p_n \in \mathbb{R}$ are coefficients specifying the function. Given data points $(u^{(1)}, y^{(1)}), \dots, (u^{(N)}, y^{(n+1)})$, find conditions under which there is a polynomial function of degree n that fits these points, i.e., conditions for existence of $p_0, p_1, \dots, p_n \in \mathbb{R}$, such that $p(u^{(i)}) = y^{(i)}$, for $i = 1, \dots, n+1$. We say that the polynomial function p *interpolates* the data points. Under the conditions that an interpolating polynomial function exists, explain how it can be computed algorithmically in terms of standard linear algebra operations. Is it unique? Check the interpolation conditions for the data

(a) $u^{(1)} = -1,\ y^{(1)} = 1, \quad u^{(2)} = 0,\ y^{(2)} = 0, \quad u^{(3)} = 1,\ y^{(3)} = 1;$ and

(b) $u^{(1)} = 0,\ y^{(1)} = 1, \quad u^{(2)} = 0,\ y^{(2)} = 0, \quad u^{(3)} = 1,\ y^{(3)} = 1$

and if they are satisfied compute a minimal degree interpolating polynomials.

Solution: We are aiming to find a polynomial function that fits the given input/output pairs. As stated in the problem formulation, "p fits the points $(u^{(i)}, y^{(i)})$, $i = 1, \dots, n+1$" means that "$p(u^{(i)}) = y^{(i)}$, for $i = 1, \dots, n+1$". Write the latter more explicitly as

$$p_0 (u^{(i)})^0 + p_1 (u^{(i)})^1 + \cdots + p_n (u^{(i)})^n = y^{(i)}, \qquad \text{for} \quad i = 1, \dots, n+1. \tag{$**$}$$

This shows that, with respect to the polynomial coefficients $p_0, p_1, \dots, p_n$, $(**)$ is a *linear* system of $n+1$ equations in $n+1$ unknowns. As we are already used to, we write it in a matrix form

$$Vp = Y,$$

where

$$V := \begin{bmatrix} (u^{(1)})^0 & (u^{(1)})^1 & \cdots & (u^{(1)})^n \\ \vdots & \vdots & & \vdots \\ u^{(n+1)} & (u^{(n+1)})^1 & \cdots & (u^{(n+1)})^n \end{bmatrix}, \qquad Y = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(n+1)} \end{bmatrix}$$

and (with some abuse of notation) $p := \operatorname{col}(p_0, p_1, \dots, p_n)$. (The matrix V is called a Vandermonde matrix)

Under the condition that V is nonsingular, there is a unique interpolating polynomial for the data, given by a vector of coefficients:

$$p = V^{-1} Y.$$

It turns out that the uniqueness condition is equivalent to $u^{(i)} \neq u^{(j)}$, for all $i \neq j$.

(a) In this example,

$$V = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \qquad \text{and} \qquad p = V^{-1} Y = \begin{bmatrix} 0 & 1 & 0 \\ -1/2 & 0 & 1/2 \\ 1/2 & -1 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Therefore, $p(u) = u^2$.

(b) In this case, the system $Vp = Y$ is incompatible, so that there is no interpolating polynomial of degree 2.

□