Lehrstuhl für STEUERUNGS- UND REGELUNGSTECHNIK Technische Universität München Prof. Dongheui Lee	MACHINE LEARNING IN ROBOTICS Exercises 1: Linear Regression

Exercise 1

Given the dataset shown in table 1 and illustrated in figure 1, we want to predict the output value for $x = 1$. We assume a linear regression model.

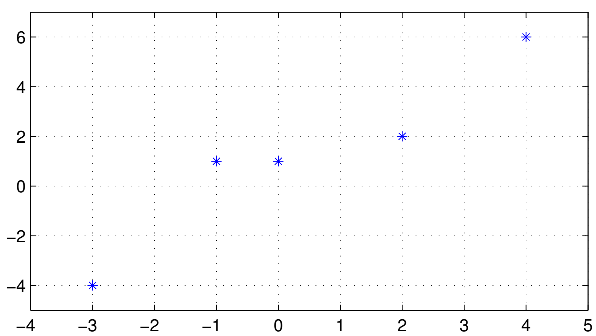

Figure 1: Training dataset

input x	-3	-1	0	2	4
output y	-4	1	1	2	6

Table 1: Data

a) Let's assume $f(x) = wx$ as a regression model with unknown parameter w. Find w which fits the data best in the sense of the Euclidean norm.

b) Let's assume $f(x) = w_0 + w_1 x$ as a regression model with unknown parameter vector $\boldsymbol{w} = [w_0 \quad w_1]^T$. By the use of the normal equation, find the best $\boldsymbol{w}$.

c) Predict the output value of the system for $x = 1$ using both regression models (a) and (b).

d) Let's assume the regression model as in (a). Now, compute the unknown parameter w by the gradient descent algorithm. Start with an initial value of $w = 0$ and use the learning rate $\alpha = 0.1$. Compute the first 2 iterations.

Solution Exercise 1

a) We have to minimize the quantity:

$$\|w\boldsymbol{x} - \boldsymbol{y}\| \Longrightarrow \|w\boldsymbol{x} - \boldsymbol{y}\|^2 = (w\boldsymbol{x} - \boldsymbol{y})^T(w\boldsymbol{x} - \boldsymbol{y}) = w^2\boldsymbol{x}^T\boldsymbol{x} - w\boldsymbol{x}^T\boldsymbol{y} - w\boldsymbol{y}^T\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{y}$$
$$= w^2\|\boldsymbol{x}\|^2 - 2w\boldsymbol{x}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2$$

considering that $\|\boldsymbol{x}\|^2 = 30$,$\|\boldsymbol{y}\|^2 = 58$ and $\boldsymbol{x}^T\boldsymbol{y} = 39$ and deriving the previous equation:

$$\frac{\partial}{\partial w}(w^2\|\boldsymbol{x}\|^2 - 2w\boldsymbol{x}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2) = 60w - 78 = 0 \Longrightarrow w^* = 1.3.$$

b) Using equations (1.6) in the lecture notes it is easy to compute the optimal parameters:

$$w_0 = \frac{n\sum x^{(i)}y^{(i)} - \sum x^{(i)}\sum y^{(i)}}{n\sum x^{(i)2} - (\sum x^{(i)})^2} = 0.6986$$
$$w_1 = \frac{\sum y^{(i)}\sum x^{(i)2} - \sum x^{(i)}\sum x^{(i)}y^{(i)}}{n\sum x^{(i)2} - (\sum x^{(i)})^2} = 1.2534.$$

otherwise:

$$\boldsymbol{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix}$$

$$\boldsymbol{w}^* = (\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y}$$

$$\boldsymbol{X^T X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -3 & -1 & 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 30 \end{bmatrix}$$

$$(\boldsymbol{X^T X})^{-1} = \frac{1}{150-4}\begin{bmatrix} 30 & -2 \\ -2 & 5 \end{bmatrix} = \frac{1}{146}\begin{bmatrix} 30 & -2 \\ -2 & 5 \end{bmatrix}$$

$$\boldsymbol{X^T y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -3 & -1 & 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ 39 \end{bmatrix}$$

$$(\boldsymbol{X^T X})^{-1}\boldsymbol{X^T y} = \frac{1}{146}\begin{bmatrix} 30 & -2 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 6 \\ 39 \end{bmatrix} = \frac{1}{146}\begin{bmatrix} 180-78 \\ -12+195 \end{bmatrix} = \frac{1}{146}\begin{bmatrix} 102 \\ 183 \end{bmatrix} = \begin{bmatrix} 0.6986 \\ 1.2534 \end{bmatrix}$$

c) $x = 1 \Longrightarrow y = 1.3$ (model (a)), $y = 1.9520$ (model (b)) .

d) $w = 0, \alpha = 0.1$:

$$w_j := w_j - \frac{\alpha}{n}\sum_{i=1}^{n}(f(\boldsymbol{x}^{(i)}) - y^{(i)})x_j^{(i)}$$

$$i = 1 \Longrightarrow w := 0 - \frac{0.1}{5}(4(-3) + (-1)(-1) + (-1)0 + (-2)2 + (-6)4)$$
$$:= -\frac{0.1}{5}(-12 + 1 - 4 - 24) := \frac{3.9}{5} := 0.78$$
$$i = 2 \Longrightarrow w := 0.78 - \frac{0.1}{5}((-3w+4)(-3) + (-w-1)(-1) + (-1)0 + (2w-2)(2)+$$
$$+ (4w-6)(4)) := 0.78 - \frac{0.1}{5}(9w + w + 4w + 16w - 12 + 1 - 4 - 24)$$
$$:= 0.78 - \frac{0.1}{5}(30w - 39) := \frac{0.1}{5}(23.4 - 39) := 1.0920.$$

Exercise 2

The kinematic of a differential-drive mobile robot like that in figure 2 is described in the discrete-time by the set of equations

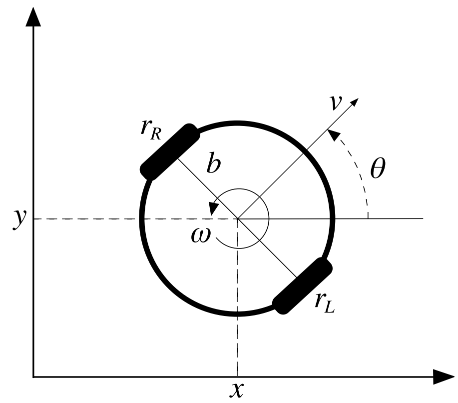

Figure 2: Top-view sketch of a differential-drive mobile robot with relevant variables.

$$\begin{cases} x^{(t+1)} = x^{(t)} + v^{(t)} cos(\theta^{(t)} + \omega^{(t)} \frac{\Delta T}{2}) \Delta T \\ y^{(t+1)} = y^{(t)} + v^{(t)} sin(\theta^{(t)} + \omega^{(t)} \frac{\Delta T}{2}) \Delta T \\ \theta^{(t+1)} = \theta^{(t)} + \omega^{(t)} \Delta T \end{cases}$$

where ΔT is the sample time. The relation between the linear v and angular ω velocities of the robot and the velocity of the wheels (ω_R and ω_L) is

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} \omega_R \\ \omega_L \end{bmatrix} = \boldsymbol{W} \begin{bmatrix} \omega_R \\ \omega_L \end{bmatrix}$$

Given m motion trajectories $T_r = \left[\left\{ x_1^{(t)}, y_1^{(t)}, \theta_1^{(t)}, \omega_{R,1}^{(t)}, \omega_{L,1}^{(t)} \right\}_{t=0}^{n}, \ldots, \left\{ x_m^{(t)}, y_m^{(t)}, \theta_m^{(t)}, \omega_{R,1}^{(t)}, \omega_{L,1}^{(t)} \right\}_{t=0}^{n} \right]$, estimate the unknown parameters $\boldsymbol{W}$ using least square regression. (Hint: $[w_{11}, w_{12}]$ and $[w_{21}, w_{22}]$ can be separately estimated.)

Solution Exercise 2

Let's rewrite the expression of $\theta^{(t+1)}$ to underline the dependence on the unknown parameters $[w_{21}, w_{22}]$. For $t = 0$ it holds that

$$\theta^{(1)} = \theta^{(0)} + w_{21} \Delta T \omega_R^{(0)} + w_{22} \Delta T \omega_L^{(0)} \ ,$$

for the final instant n it holds that

$$\theta^{(n)} = \theta^{(0)} + w_{21} \Delta T \sum_{t=0}^{n-1} \omega_R^{(t)} + w_{22} \Delta T \sum_{t=0}^{n-1} \omega_L^{(t)} \ ,$$

that can be written in a compact form by choosing $\boldsymbol{X}_\theta = \Delta T \left[\sum_{t=0}^{n-1} \omega_R^{(t)} \ \sum_{t=0}^{n-1} \omega_L^{(t)} \right]$

$$\theta^{(n)} - \theta^{(0)} = \boldsymbol{X}_\theta \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix}$$

The m given trajectories can be stacked in the form

$$\begin{bmatrix} \theta_1^{(n)} - \theta_1^{(0)} \\ \vdots \\ \theta_m^{(n)} - \theta_m^{(0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{\theta,1} \\ \vdots \\ \boldsymbol{X}_{\theta,m} \end{bmatrix} \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix} = \bar{\boldsymbol{X}}_\theta \begin{bmatrix} w_{21} \\ w_{22} \end{bmatrix}$$

and the optimal values for $[w_{21}, w_{22}]$ estimated as

$$\begin{bmatrix} w_{21}^* \\ w_{22}^* \end{bmatrix} = \left(\bar{\boldsymbol{X}}_\theta^T \bar{\boldsymbol{X}}_\theta\right)^{-1} \bar{\boldsymbol{X}}_\theta^T \begin{bmatrix} \theta_1^{(n)} - \theta_1^{(0)} \\ \vdots \\ \theta_m^{(n)} - \theta_m^{(0)} \end{bmatrix} .$$

Following a similar reasoning it is possible to estimate $[w_{11}, w_{12}]$. Let's rewrite the expression of $x^{(t+1)}$ and $y^{(t+1)}$ for $t = 0$. To easy the notation, we substitute $\alpha^{(i)} = \theta^{(i)} + \omega^{(i)} \frac{\Delta T}{2}$.

$$\begin{cases} x^{(1)} = x^{(0)} + w_{11} \Delta T \omega_R^{(0)} cos(\alpha^{(0)}) + w_{21} \Delta T \omega_L^{(0)} cos(\alpha^{(0)}) \\ y^{(1)} = y^{(0)} + w_{11} \Delta T \omega_R^{(0)} sin(\alpha^{(0)}) + w_{21} \Delta T \omega_L^{(0)} sin(\alpha^{(0)}) \end{cases}$$

For the final instant n it holds that

$$\begin{cases} x^{(n)} - x^{(0)} = w_{11} \Delta T \sum_{t=0}^{n-1} \omega_R^{(t)} cos(\alpha^{(t)}) + w_{21} \Delta T \sum_{t=0}^{n-1} \omega_L^{(t)} cos(\alpha^{(t)}) \\ y^{(n)} - y^{(0)} = w_{11} \Delta T \sum_{t=0}^{n-1} \omega_R^{(t)} sin(\alpha^{(t)}) + w_{21} \Delta T \sum_{t=0}^{n-1} \omega_L^{(t)} sin(\alpha^{(t)}) \end{cases}$$

that can be written in the compact form

$$\begin{bmatrix} x^{(n)} - x^{(0)} \\ y^{(n)} - y^{(0)} \end{bmatrix} = \boldsymbol{X}_{xy} \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix} .$$

by choosing

$$\boldsymbol{X}_{xy} = \Delta T \begin{bmatrix} \sum_{t=0}^{n-1} \omega_R^{(t)} cos(\alpha^{(t)}) & \sum_{t=0}^{n-1} \omega_L^{(t)} cos(\alpha^{(t)}) \\ \sum_{t=0}^{n-1} \omega_R^{(t)} sin(\alpha^{(t)}) & \sum_{t=0}^{n-1} \omega_L^{(t)} sin(\alpha^{(t)}) \end{bmatrix} .$$

The m given trajectories can be stacked in the form

$$\begin{bmatrix} x_1^{(n)} - x_1^{(0)} \\ y_1^{(n)} - y_1^{(0)} \\ \vdots \\ x_m^{(n)} - x_m^{(0)} \\ y_m^{(n)} - y_m^{(0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{xy,1} \\ \vdots \\ \boldsymbol{X}_{xy,m} \end{bmatrix} \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix} = \bar{\boldsymbol{X}}_{xy} \begin{bmatrix} w_{11} \\ w_{12} \end{bmatrix}$$

and the optimal values for $[w_{11}, w_{12}]$ estimated as

$$\begin{bmatrix} w_{11}^* \\ w_{12}^* \end{bmatrix} = \left(\bar{\boldsymbol{X}}_{xy}^T \bar{\boldsymbol{X}}_{xy}\right)^{-1} \bar{\boldsymbol{X}}_{xy}^T \begin{bmatrix} x_1^{(n)} - x_1^{(0)} \\ y_1^{(n)} - y_1^{(0)} \\ \vdots \\ x_m^{(n)} - x_m^{(0)} \\ y_m^{(n)} - y_m^{(0)} \end{bmatrix} .$$

Exercise 3

Given the dataset in Exercise 1, we want to predict the output value for $x = 1$ using a quadratic regression model.

a) Let's assume $f(x) = w_1 x + w_2 x^2$ as a regression model with unknown parameter vector $\boldsymbol{w} = [w_1 \quad w_2]^T$. Find $\boldsymbol{w}$ which fits the data best in the sense of the Euclidean norm.

c) Predict the output value of the system for $x = 1$ using the regression model (a).

Solution Exercise 3

a)

$$\boldsymbol{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \ \boldsymbol{X} = \begin{bmatrix} -3 & 9 \\ -1 & 1 \\ 0 & 0 \\ 2 & 4 \\ 4 & 16 \end{bmatrix}, \ \boldsymbol{y} = \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix}$$

$$\boldsymbol{w}^* = (\boldsymbol{X^T X})^{-1} \boldsymbol{X^T y}$$

$$\boldsymbol{X^T X} = \begin{bmatrix} -3 & -1 & 0 & 2 & 4 \\ 9 & 1 & 0 & 4 & 16 \end{bmatrix} \begin{bmatrix} -3 & 9 \\ -1 & 1 \\ 0 & 0 \\ 2 & 4 \\ 4 & 16 \end{bmatrix} = \begin{bmatrix} 30 & 44 \\ 44 & 354 \end{bmatrix}$$

$$(\boldsymbol{X^T X})^{-1} = \frac{1}{10620 - 1936} \begin{bmatrix} 354 & -44 \\ -44 & 30 \end{bmatrix} = \frac{1}{8684} \begin{bmatrix} 354 & -44 \\ -44 & 30 \end{bmatrix}$$

$$\boldsymbol{X^T y} = \begin{bmatrix} -3 & -1 & 0 & 2 & 4 \\ 9 & 1 & 0 & 4 & 16 \end{bmatrix} \begin{bmatrix} -4 \\ 1 \\ 1 \\ 2 \\ 6 \end{bmatrix} = \begin{bmatrix} 39 \\ 69 \end{bmatrix}$$

$$(\boldsymbol{X^T X})^{-1} \boldsymbol{X^T y} = \frac{1}{8684} \begin{bmatrix} 354 & -44 \\ -44 & 30 \end{bmatrix} \begin{bmatrix} 39 \\ 69 \end{bmatrix} = \frac{1}{8684} \begin{bmatrix} 13806 - 3036 \\ -1716 + 2070 \end{bmatrix}$$

$$= \frac{1}{8684} \begin{bmatrix} 10770 \\ 354 \end{bmatrix} = \begin{bmatrix} 1.2402 \\ 0.0408 \end{bmatrix}$$

b) $x = 1 \Longrightarrow y = 1.2402 * 1 + 0.0408 * 1^2 = 1.281$.