Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Dr. Francis Colas
Institute of Robotics and Intelligent Systems
Autonomous Systems Lab

ETH Zürich
CLA E 26
Tannenstraße 3
8092 Zürich Switzerland

fcolas@mavt.ethz.ch
www.asl.ethz.ch

Information Processing in Robotics
Solution Sheet 4
Topics: Regression

Solution to exercise 1:

(a) In linear regression, the model looked for is a linear combination of the basis functions, and this yields a set of linear equations. Regression can be non-linear if the parameters are not just relative weights, like for example when fitting a sine function $A \sin(\omega t + \theta)$. The equations from the minimization constraint are not linear anymore and other techniques are needed.

(b) The trajectory should be straight in the (x, y) plane and the polynomial basis should be good. Note that x (or y) as a function of t are not expected to be linear due to friction, for example.

(c)

$$E(w) = \frac{1}{2}\sum_t (w\phi(x_t) - y_t)^2 \quad (1)$$

$$E(w) = \frac{1}{2}\sum_t (w - y_t)^2 \quad (2)$$

$\hat{w} = argmin_w E(w)$ is such that $\frac{dE}{dw}(\hat{w}) = 0$:

$$\frac{dE}{dw}(\hat{w}) = 0 \quad (3)$$

$$\frac{1}{2}\sum_t 2.(\hat{w} - y_t) = 0 \quad (4)$$

$$N.\hat{w} - \sum_t y_t = 0 \quad (5)$$

$$\hat{w} = \frac{1}{N}\sum_t y_t \quad (6)$$

A polynomial fit of order 0 is nothing else than the mean of the points.

Solution to exercice 2:

(a)

$$E(a,b) = \frac{1}{2}\sum_t (a.x_t + b - y_t)^2$$

(b)

$$\begin{cases} \frac{\partial E}{\partial a}(a,b) &= \sum_t x_t\,(a.x_t + b - y_t) \\ \frac{\partial E}{\partial b}(a,b) &= \sum_t (a.x_t + b - y_t) \end{cases}$$

Minimizing the error means finding places where its gradient is null; that is, each partial derivative is zero:

$$\begin{cases} \sum_t x_t\,(a.x_t + b - y_t) &= 0 \\ \sum_t (a.x_t + b - y_t) &= 0 \end{cases}$$

(c) The second equation can be rewritten:

$$\sum_t (a.x_t + b - y_t) = 0 \quad (7)$$

$$a.\sum_t x_t + \sum_t b - \sum_t y_t = 0 \quad (8)$$

$$a.\bar{x} + b - \bar{y} = 0 \quad (9)$$

That is: $(\bar{x}, \bar{y})$ lies necessarily on the fitted line. From that we have:

$$b = \bar{y} - a.\bar{x}$$

(d) With this expression for b we get:

$$\begin{cases} \sum_t x_t\,(a.x_t + b - y_t) &= 0 \\ \sum_t (a.x_t + b - y_t) &= 0 \end{cases} \quad (10)$$

$$\begin{cases} \sum_t x_t\,(a.x_t + (\bar{y} - a.\bar{x}) - y_t) &= 0 \\ \sum_t (a.x_t + (\bar{y} - a.\bar{x}) - y_t) &= 0 \end{cases} \quad (11)$$

$$\begin{cases} \sum_t x_t\,(a(x_t - \bar{x}) - (y_t - \bar{y})) &= 0 \\ \sum_t (a(x_t - \bar{x}) - (y_t - \bar{y})) &= 0 \end{cases} \quad (12)$$

Now if we multiply the second equation by $\bar{x}$ and substract it to the first equation, we get:

$$\sum_t (x_t - \bar{x})\,(a(x_t - \bar{x}) - (y_t - \bar{y})) = 0 \quad (13)$$

$$a.\sum_t (x_t - \bar{x})^2 - \sum_t (x_t - \bar{x})\,(y_t - \bar{y}) \quad 0 \quad (14)$$

$$a = \frac{\sum_t (x_t - \bar{x})\,(y_t - \bar{y})}{\sum_t (x_t - \bar{x})^2} \quad (15)$$

That is, a is the ratio between the (empirical) covariance of x_t and y_t, and the (empirical) variance on x_t.

(e) $a \approx -3.9427$ and $b \approx 3.9777$

We want the position of the intersection between $y = a.x + b$ and the goal line $y = 0$: $x = \frac{-b}{a} \approx 1.0089$. The ball should pass less than 9mm out of the goal: it should be ok but depending on how precise we trust our sensors and the model (notably its linear assumption) we may want to try and block the ball.

(f) By symmetry we have:

$$\alpha = \frac{\sum_t (y_t - \bar{y})(x_t - \bar{x})}{\sum_t (y_t - \bar{y})^2} \approx -0.2531$$

and

$$\beta = \bar{x} - \alpha.\bar{y} \approx 1.0071$$

Now the ball should cross the line on $(1.0071, 0)$ (compared to $(1.0089, 0)$ above).

(g) Substituting $y = a'.x + b'$ in $x = \alpha.y + \beta$ and solving for a' and b' we have:

$$x = \alpha.(a'.x + b') + \beta \tag{16}$$

$$\iff x = \alpha.a'.x + \alpha.b' + \beta \tag{17}$$

$$\iff \begin{cases} 1 = \alpha.a' \\ 0 = \alpha.b' + \beta \end{cases} \tag{18}$$

$$\iff \begin{cases} a' = \frac{1}{\alpha} \\ b' = -\frac{\beta}{\alpha} \end{cases} \tag{19}$$

Numerically we have: $a' \approx -3.9505 \neq -3.9427 \approx a$
and $b' \approx 3.9787 \neq 3.9777 \approx b$.

$$E(a, b) = \frac{1}{2} \sum_t (a.x_t + b - y_t)^2 \approx 9.666e - 3$$

$$E(\alpha, \beta) \approx 9.685e - 3$$

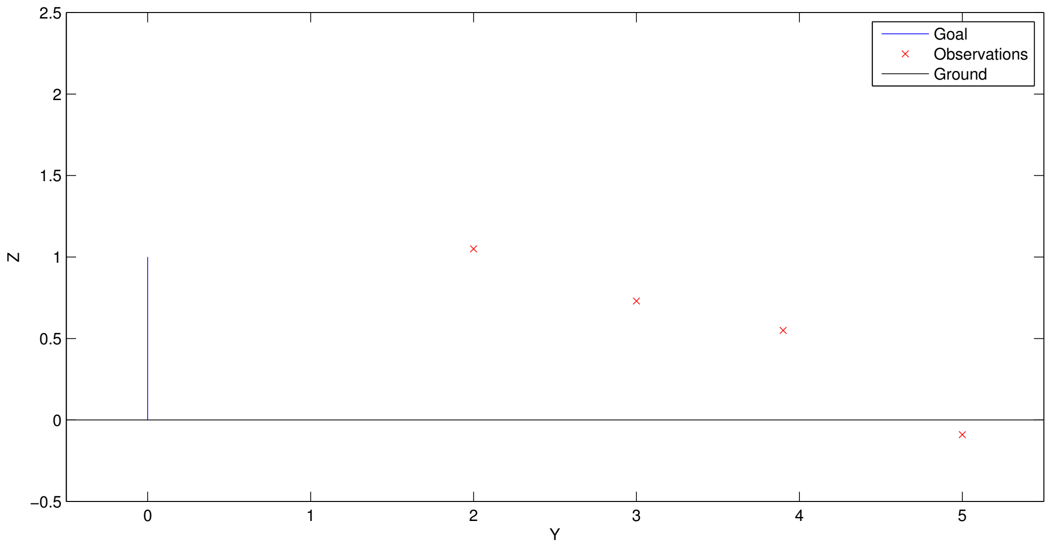
The two lines are different and the first one is indeed the least squared error with respect to a regression of y against x. You can likewise say that it is the reverse for x against y.

The reason for it is that the error used is not the shortest distance to the line but the distance to the point with the same x-coordinate: this is not symmetric in x

and y. The underlying assumption is that the x values are correct and only the y values are noisy. With the regression of x against y it is the contrary: x is assumed to be noisy whereas y is not.

In our case, as both x and y are measured, neither assumptions hold and a error defined as the distance to the line would be preferable. Note that such error function would be difficult to generalize to arbitrary curves for which the distance to a point can be costly to compute.

Solution to exercise 3:



(a)

(b) Order 1: $\Phi_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}$ Order 2: $\Phi_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix}$

Order 3: $\Phi_3 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix}$

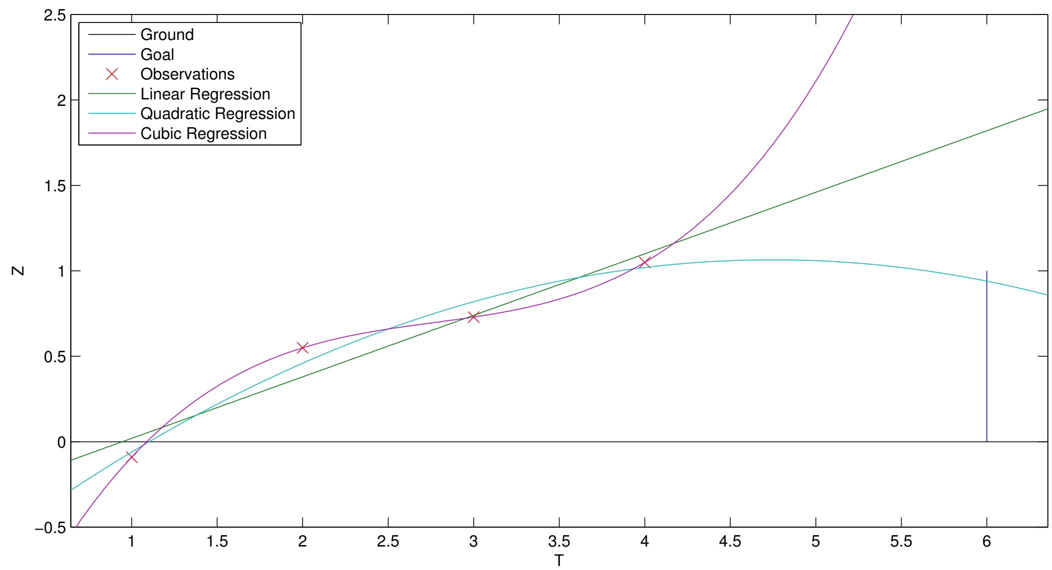
We use $\vec{w} = \Phi^+ \times Z^T$ where Φ^+ is the pseudo-inverse of Φ and Z^T is the column vector of the altitude observations.

We have: $\vec{w}_1 \approx \begin{pmatrix} -0.34 \\ 0.36 \end{pmatrix}$, $\vec{w}_2 \approx \begin{pmatrix} -0.74 \\ 0.76 \\ -0.08 \end{pmatrix}$, and $\vec{w}_3 \approx \begin{pmatrix} -1.79 \\ 2.43 \\ -0.83 \\ 0.1 \end{pmatrix}$.

(c) Prediction at frame 6:

- order 1: ≈ 1.82,
- order 2: ≈ 0.94,
- order 3: ≈ 4.51.

Only order 2 regression claims that the ball enters the goal (with respect to its altitude).



Looking at the trajectories predicted by each regression, it seems that order 3 is grossly overfitting our data and we should discard it. From daily experience we don't expect the ball to continue getting up indefinitely, as predicted by order 1. Moreover, with a bit more physics expertise, we expect the gravity to shape the trajectory as a parabola, as proposed by order 2 regression.[1] As a consequence, we should probably trust the order 2 regression.

[1] From this observation and the order 2 coefficient in the fit, we can estimate the framerate of our observations to around 11Hz.