MATH 115 LECTURE 22: Least-Squares and Markov chains.

Last day: (..., pages -158-, -159-)

• Exercise: Consider the data points $x = 0, 1, 2$, $y = 0, 1, 5$. We want to fit a model $y = ax^2 + bx$ that minimizes the squares error.

Sol: Errors: $e_1 = y_1 - (ax_1^2 + bx_1) = 0 - 0 = 0$

$e_2 = y_2 - (ax_2^2 + bx_2) = 1 - a - b$

$e_3 = y_3 - (ax_3^2 + bx_3) = 5 - 4a - 2b$

Total error (square): $E(a,b) = e_1^2 + e_2^2 + e_3^2 =$

$= (1-a-b)^2 + (5-4a-2b)^2$

$\min_{a,b} E(a,b) \rightsquigarrow \nabla E(a,b) = (0,0) \rightsquigarrow$

$$\left.\begin{aligned} \frac{\partial E}{\partial a}(a,b) &= -2(1-a-b) - 8(5-4a-2b) = 0 \\ \frac{\partial E}{\partial b}(a,b) &= -2(1-a-b) - 4(5-4a-2b) = 0 \end{aligned}\right\} \rightarrow$$

$$\rightarrow \left| \begin{array}{l} (2+32)a + (2+16)b = 2+40 \\ (2+18)a + (2+8)b = 2+20 \end{array} \right\} \rightarrow \begin{array}{l} 34a + 18b = 42 \\ 18a + 10b = 22 \end{array} \Big\} \rightarrow$$

$$\rightarrow \left[\begin{array}{cc|c} 34 & 18 & 42 \\ 18 & 10 & 22 \end{array} \right] \xrightarrow{R2' = R2 - \frac{18}{34} R1} \left[\begin{array}{cc|c} 34 & 18 & 42 \\ 0 & \underbrace{10 - \frac{18}{34} 18}_{= \frac{16}{34}} & \underbrace{22 - \frac{18}{34} 42}_{= \frac{-8}{34}} \end{array} \right] \rightarrow$$

$$\Rightarrow b = \frac{34}{16} \frac{-8}{34} = \frac{-1}{2}$$

$$\hookrightarrow 34a = 42 - 18 \frac{-1}{2} \Rightarrow a = \frac{51}{34}.$$

So the model is $y = \frac{51}{34} x^2 - \frac{1}{2} x$.

<u>Remark</u>: Notice that in this example the model perfectly fits the data. That is, the total error is zero!

$\hookrightarrow$ We could have noticed this from the beginning:

$$\left. \begin{array}{l} (0,0) \rightarrow 0 = 0 \\ (1,1) \rightarrow 1 = a + b \\ (2,5) \rightarrow 5 = 4a + 2b \end{array} \right\} \text{This system has a solution!}$$

$\rightarrow$ In general, if a system of equations has a solution, then the least squares solution is that same solution.

• <u>Exercise:</u> Fit the model $y = a\sin(t) + b\,t\cos(t)$ using the least squares method, given the data points $\begin{cases} t = 0, \frac{-\pi}{2}, \frac{\pi}{2}, \pi, \\ y = 0, 1, \frac{-1}{2}, \frac{-\pi}{2} \end{cases}$

<u>Sol</u>

Errors: $e_1 = y_1 - (a\sin(t_1) + b t_1 \cos(t_1)) = 0 - 0 = 0.$

$$e_2 = 1 - \left(a\sin\left(\tfrac{-\pi}{2}\right) - \tfrac{\pi}{2} b \cos\left(\tfrac{-\pi}{2}\right)\right) = 1 + a$$

$$e_3 = \frac{-1}{2} - \left(a\sin\left(\tfrac{\pi}{2}\right) + b\tfrac{\pi}{2}\cos\left(\tfrac{\pi}{2}\right)\right) = \frac{-1}{2} - a$$

$$e_4 = \frac{-\pi}{2} - \left(a\sin(\pi) + b\pi\cos(\pi)\right) = \frac{-\pi}{2} + \pi b.$$

Total squares error:

$$E(a,b) = (1+a)^2 + \left(\tfrac{1}{2} + a\right)^2 + \left(\tfrac{-\pi}{2} + \pi b\right)^2.$$

Minimize $E(a,b)$:

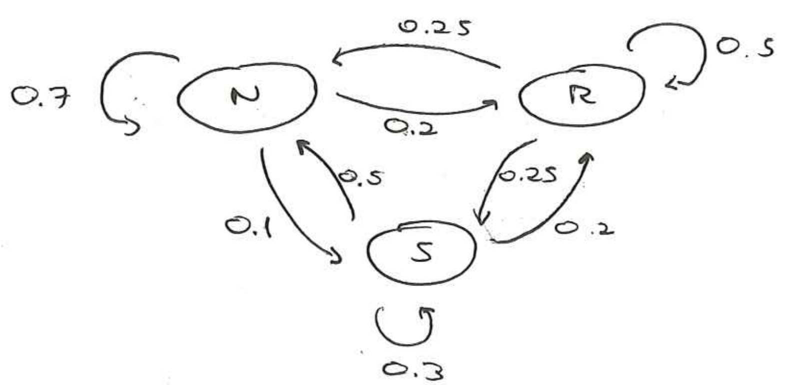
$$\frac{\partial E}{\partial a}(a,b) = 2(1+a) + 2\left(\tfrac{1}{2} + a\right) = 0 \Leftrightarrow 4a = -3 \Rightarrow a = \frac{-3}{4}$$

$$\frac{\partial E}{\partial b}(a,b) = 2\pi\left(\tfrac{-\pi}{2} + \pi b\right) = 0 \Leftrightarrow 2\pi^2 b = \pi^2 \Leftrightarrow b = \frac{1}{2}.$$

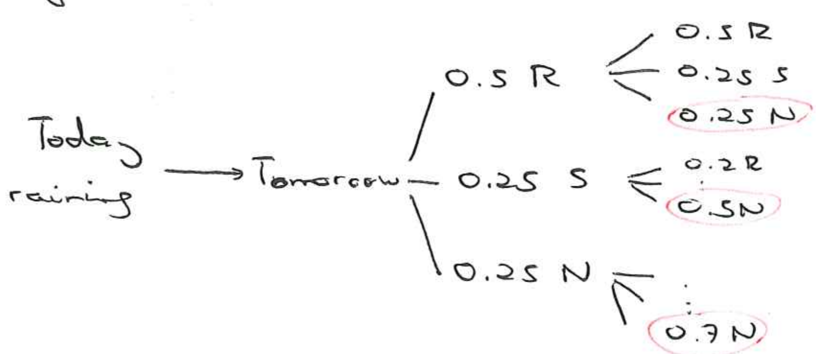
So the model is $y = \frac{-3}{4}\sin(t) + \frac{1}{2}t\cos(t).$

New Chapter: Markov Chains.

Consider the following silly model for weather prediction:



If today is raining, what is the probability that the day after tomorrow is a nice day?



So after tomorrow,

$$p = (0.5)(0.25) + (0.25)(0.5) + (0.25)(0.7) = 0.425$$

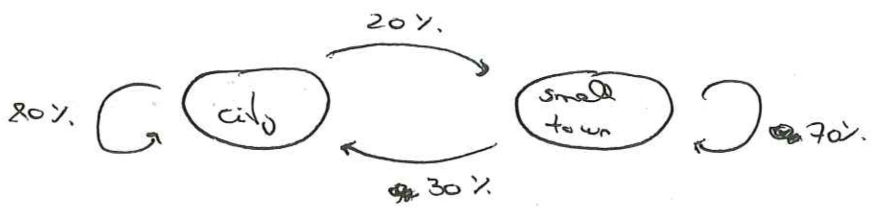
What is the probability that it will snow in a week, if today is a nice day?

Moreover, along the years, which fraction of days are nice, rainy or snowy?

→ The answers are not straightforward, but we will see they are easy to compute.

• Let's use a different simpler example:

(step = 1 year)



If today 50% of people live in cities, what is the distribution next year?

In cities: $(0.5)\cdot(0.8) + (0.5)(0.3) = 0.55 \to 55\%$.

In towns: $(0.5)\cdot(0.7) + (0.5)\cdot(0.2) = 0.45 \to 45\%$.

(same numbers)

And next one:

In cities: $(0.55)\cdot(0.8) + (0.45\cdot 0.3) = 0.575 \to 57'5\%$.

So in towns → 42'5%.

Create the "state" vector $\vec{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$, where

$x_1(k) = \%$ in cities in year k
$x_2(k) = \%$ in town in year k

(of course, $x_1(k) + x_2(k) = 1 \quad \forall k$).

How to find $\vec{x}(k+1)$?

$$\left.\begin{aligned} x_1(k+1) &= x_1(k)\cdot(0.8) + x_2(k)\cdot(0.3) \\ x_2(k+1) &= x_1(k)\cdot(0.2) + x_2(k)(0.7) \end{aligned}\right\} \rightarrow \vec{x}(k+1) = \underbrace{\begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}}_{\text{transition matrix}} \vec{x}(k)$$

Notice: columns add up to 1.

Q: What is the transition matrix for the weather model?

State vector: $\vec{x} = \begin{bmatrix} x_S \\ x_N \\ x_R \end{bmatrix}$,

$$\vec{x}(k+1) = P\,\vec{x}(k) \rightarrow P = \begin{bmatrix} 0.3 & 0.1 & 0.25 \\ 0.5 & 0.7 & 0.25 \\ 0.2 & 0.2 & 0.5 \end{bmatrix}$$

(P_{ij} = probability of going in one step from i to j)

Check previous result:

$$\underset{\substack{\uparrow \\ \text{probabilities} \\ \text{after two} \\ \text{days}}}{p} = P^2 \underset{\substack{\uparrow \\ \text{today} \\ \text{raining}}}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} = P \begin{bmatrix} 0.25 \\ 0.25 \\ 0.5 \end{bmatrix} = \begin{bmatrix} (0.3)(0.25) + (0.1)\cdot(0.25) + (0.5)\cdot(0.25) \\ (0.5)\cdot(0.25) + (0.7)\cdot(0.25) + (0.5)(0.25) \\ \cdots \end{bmatrix}$$

$\leftarrow$ probability of nice day after 2 days.

• Def: Markov (or stochastic) matrix

An $n \times n$ matrix is Markov if:

1) All columns add up to 1.

2) All entries are ≥ 0.

• The sequence of states $\{\vec{x}(1), \vec{x}(2), \dots, \vec{x}(k), \dots\}$ defined by a ~~m~~ Markov matrix is called a Markov chain (discrete)

Remark: key feature of this model: Every state only depends on the previous step: $\vec{x}(k+1) = A\vec{x}(k)$.

Example: 1) The evolution of a particle, defined through its position and velocity, is a Markov process (by Newton's law, if we know position and velocity now, we know the future).

ok, I'm considering classical mechanics.

2) Poker is not a Markov process: you use (or should use) the knowledge about which cards have been appearing from the start.