Learning to learn

Ivan Markovsky

Problem solving is the most characteristic human activity

solved problem is an achievement of intelligence

our survival depends on solving grand problems

today's knowledge may be irrelevant tomorrow

> *"In a fast changing world, the most valuable skill is learning to learn."*
>
> *R. Hamming*

Plan

Information vs understanding

What should I understand?

Example: exploration in data modeling

Plan

Information vs understanding

What should I understand?

Example: exploration in data modeling

We learn by imitation and practice

observe common patterns (get information)

practice by using them (gain understanding)

knowledge $=$ information $+$ understanding

Information grows at exponential rate

Moore’s law applies to information

the challenge becomes to select information

we need "good" taste to navigate us

We have more information but less time to think

Google-knowing: find solution on the internet

the readily available solutions make us lazy

in order to understand, we need to discover

Classes feel boring when you are passive and not challenged

> *"College is a place where a professor's lecture notes go straight to the students' lecture notes, without passing through the brains of either."* *M. Twain*

being active is not enough: you need challenge

exams are challenging, and certainly not boring

Plan

Information vs understanding

What should I understand?

Example: exploration in data modeling

The process of solving problems is more important than the solutions

try to figure out something you don't know

the learning happens while trying to understand

vice verse, you don't learn by getting solution

. . . without first trying

No pain, no gain!

Educational system needs revision

education is virtually the same as 100 years ago

schools resemble factories

- pupils enter into batches according to age
- they follow predefined classes (like on conveyor line)
- the ones who exit pass quality specification (exams)

exams measure information, not understanding

Find your passion and follow it

the goal is self-realization, not passing exams

the motivation should be internal, not external

finding/achieving personal goals is up to you

Famous people had vision and pursued it by all means

read the biographies of those you admire

they "did not fit", but new what they wanted

they self-taught what they needed to know

Choose what and how to learn

focus on what makes you genuinely curious

ask questions; challenge established beliefs

accept nothing for granted

Plan

Information vs understanding

What should I understand?

Example: exploration in data modeling

Go from simple to complex

simple question:

check if given points lay on given line

more difficult question:

check if given points lay on a line

challenge: the line isn't given; there are ∞ many

Line fitting leads to a rank constraint

the points $(x_1, y_1), \ldots, (x_N, y_N)$ lie on a line $\mathscr{B}$

$$\Updownarrow$$

there is $(a, b) \neq 0$, such that $ax_i + by_i = 0$, for all i

$$\Updownarrow$$

there is $(a, b) \neq 0$, s.t. $[a \;\; b] \underbrace{\begin{bmatrix} x_1 & \cdots & x_N \\ y_1 & \cdots & y_N \end{bmatrix}}_{D} = 0$

$$\Updownarrow$$

$$\text{rank}(D) \leq 1$$

Reduce a new problem to an old one

we transformed one problem into another

> *Mathematics is about saying the same thing another way.*

however, rank computation is a solved problem

in this sense, we solved the original problem

Compare different solutions

shortest / most insightful

most practical: leads to algorithms

has the highest potential for generalization

The rank condition is generalizable to other data modeling problems

higher-dimensional data (subspace fitting)

time series fitting (dynamical models)

curve fitting (nonlinear models)

Approximate line fitting (noisy data) $\rightsquigarrow$ low-rank approximation

find the "nearest" matrix $\widehat{D}$ to the data matrix D,
so that there is exact fitting line for $\widehat{D}$

"nearest" means "smallest distance $\|D - \widehat{D}\|$"

mathematical problem:

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{D} \quad \|D - \widehat{D}\| \\ &\text{subject to} \quad \operatorname{rank}(\widehat{D}) \leq 1 \end{aligned}$$

High-dimensional data generalization $\rightsquigarrow$ low-rank approximation

find the "nearest" matrix $\widehat{D}$ to the data matrix D,

so that $\widehat{D} \subset \mathscr{B}$ — subspace of dimension $\leq r$

mathematical problem:

$$\begin{array}{lll} \text{minimize} & \text{over } \widehat{D} & \|D - \widehat{D}\| \\ \text{subject to} & \text{rank}(\widehat{D}) \leq r & \end{array} \qquad \text{(LRA)}$$

conjecture: (LRA) is general data modeling tool

Summary

find what interests you and be genuinely curious

learn by trying to solve challenging problems

keep trying; its a life-time long process

References

books

- G. Polya, How to solve it?
- R. Feynman, Surely you are joking, Mr. Feynman!
- R. Hamming, The art of doing science and engineering: Learning to learn

videos

- K. Robinson, Do schools kill creativity?
- E. Mazur, Confessions of a converted lecturer