

BIOMENG 261

TISSUE AND BIOMOLECULAR ENGINEERING

Oliver Maclaren
oliver.maclaren@auckland.ac.nz

MODULE OVERVIEW

Reaction kinetics and systems biology (*Oliver Maclaren*)
[11 lectures/3 tutorials/2 labs]

1. Basic principles: modelling with reaction kinetics [4 lectures]

Conservation, directional and constitutive principles. Mass action. Enzyme kinetics. Enzyme regulation. Mathematical/graphical tools for analysis and fitting.

2. Systems biology I: signalling and metabolic systems [2 lectures]

Overview of systems biology. Modelling signalling systems using reaction kinetics. Introduction to parameter estimation. Modelling metabolic systems using reaction kinetics. Flux balance analysis and constraint-based methods.

3. Systems biology II: genetic systems [3 lectures]

Modelling genes and gene regulation using reaction kinetics. Gene regulatory networks, transcriptomics and analysis of microarray data.

LECTURE 6 INTRO TO PARAMETER ESTIMATION

- Forward vs inverse problems
- Trade-offs:
 - Fit vs complexity
 - Overfitting vs underfitting
 - Training vs test
- Distances and optimal trade-off curves
- Examples

FORWARD AND INVERSE PROBLEMS

Typical modelling deals with (so-called) *forward problems*:

*Given parameters and initial conditions,
predict data*

FORWARD AND INVERSE PROBLEMS

However, in science and engineering we are usually confronted with *inverse problems*:

Given data, estimate parameters and/or initial conditions and predict future data

ILL-POSED PROBLEMS

In contrast to forward problems, inverse problems can have

- No solution
- Many solutions
- Unique but unstable solutions

These are called *ill-posed* (c.f. 'well-posed') problems
(Hadamard, 1902)

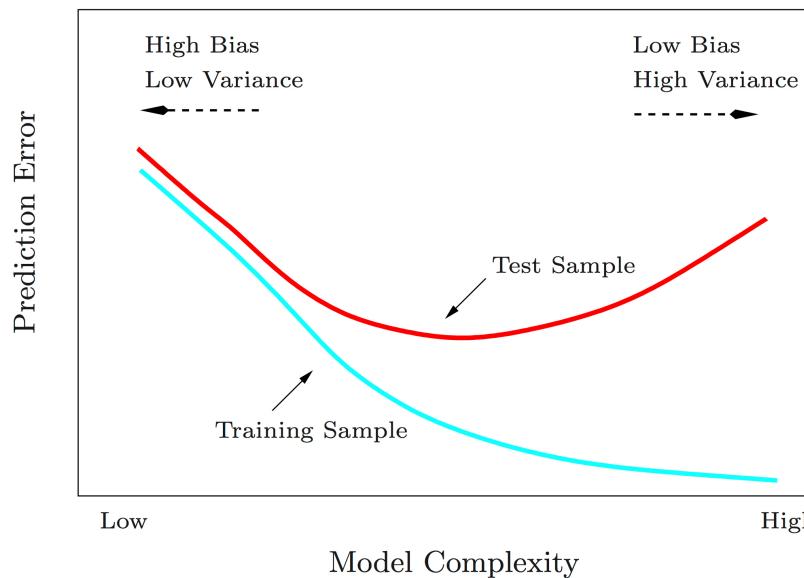
TRADE-OFFS

The key to solving ill-posed problems is recognising the
trade-offs involved

- Fit vs complexity
- Overfitting vs underfitting
- Training vs test
- Efficiency vs stability
- Bias vs variance
- Etc

These trade-offs are closely related and *recur throughout science, statistics and engineering*

TRADE-OFFS



From Hastie et. al 'Elements of Statistical Learning: Data Mining, Inference and Prediction':

available at: <http://web.stanford.edu/~hastie/ElemStatLearn/>. See also: 'An Introduction to Statistical Learning' (simplified version of above), available at: <http://www-bcf.usc.edu/~gareth/ISL/>

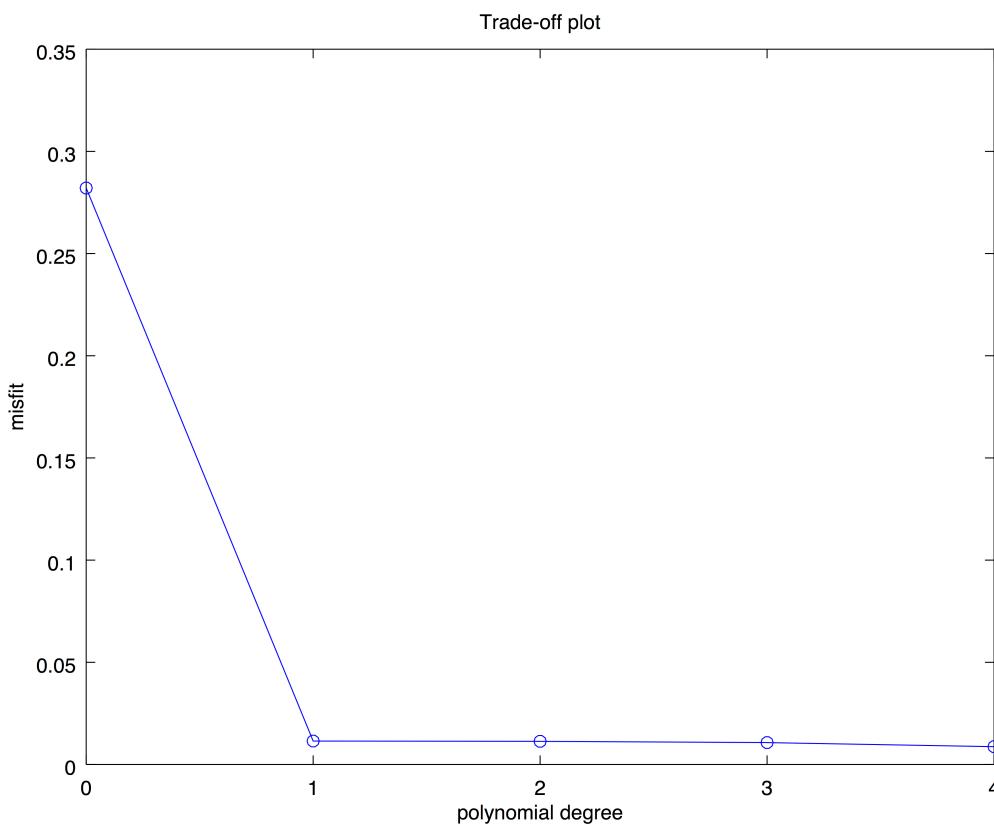
EXAMPLE: SIMPLE POLYNOMIAL FITTING

Suppose we have enzyme data and we *want to fit a curve to
the double-reciprocal plot*

- Usually use linear (first order) regression (relates to MM)

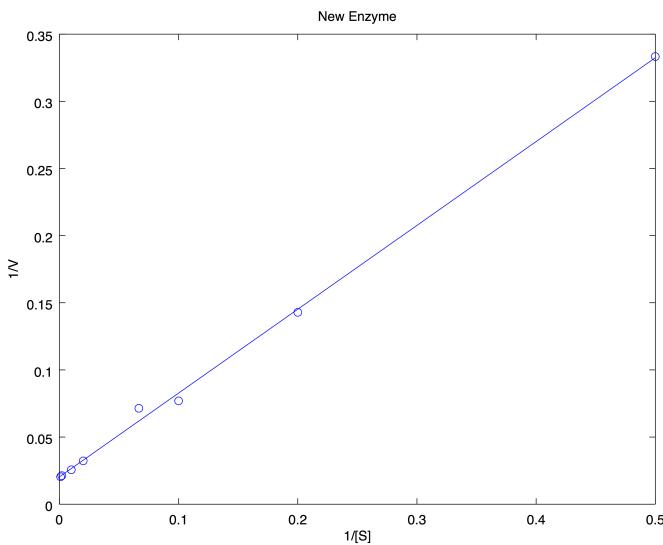
*What if we tried to fit higher order
polynomials?*

SIMPLE POLYNOMIAL FITTING: FIT VS ORDER OF POLYNOMIAL

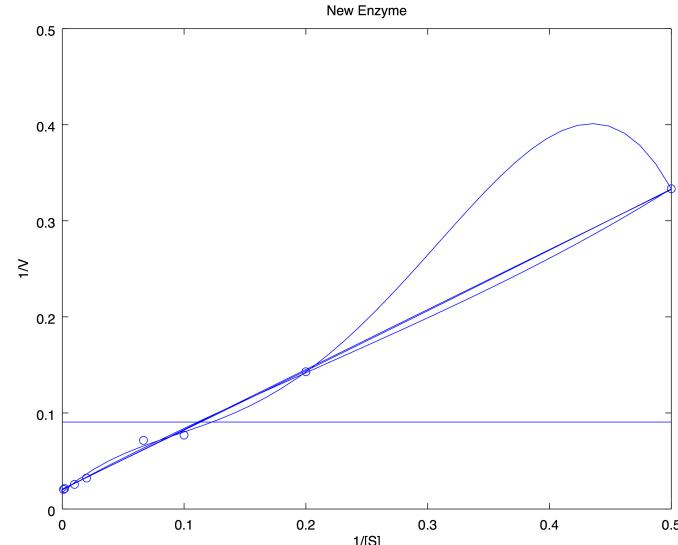


SIMPLE POLYNOMIAL FITTING: FIT VS STABILITY

Fit line

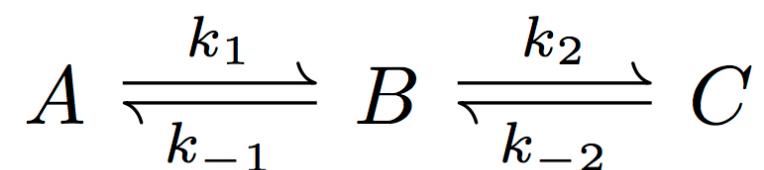


Fit polynomials



EXAMPLE: SIMPLE REACTION

Suppose we have the reaction:



Exercise: write down the differential equations (assuming mass action)!

SIMPLE REACTION: SYNTHETIC DATA

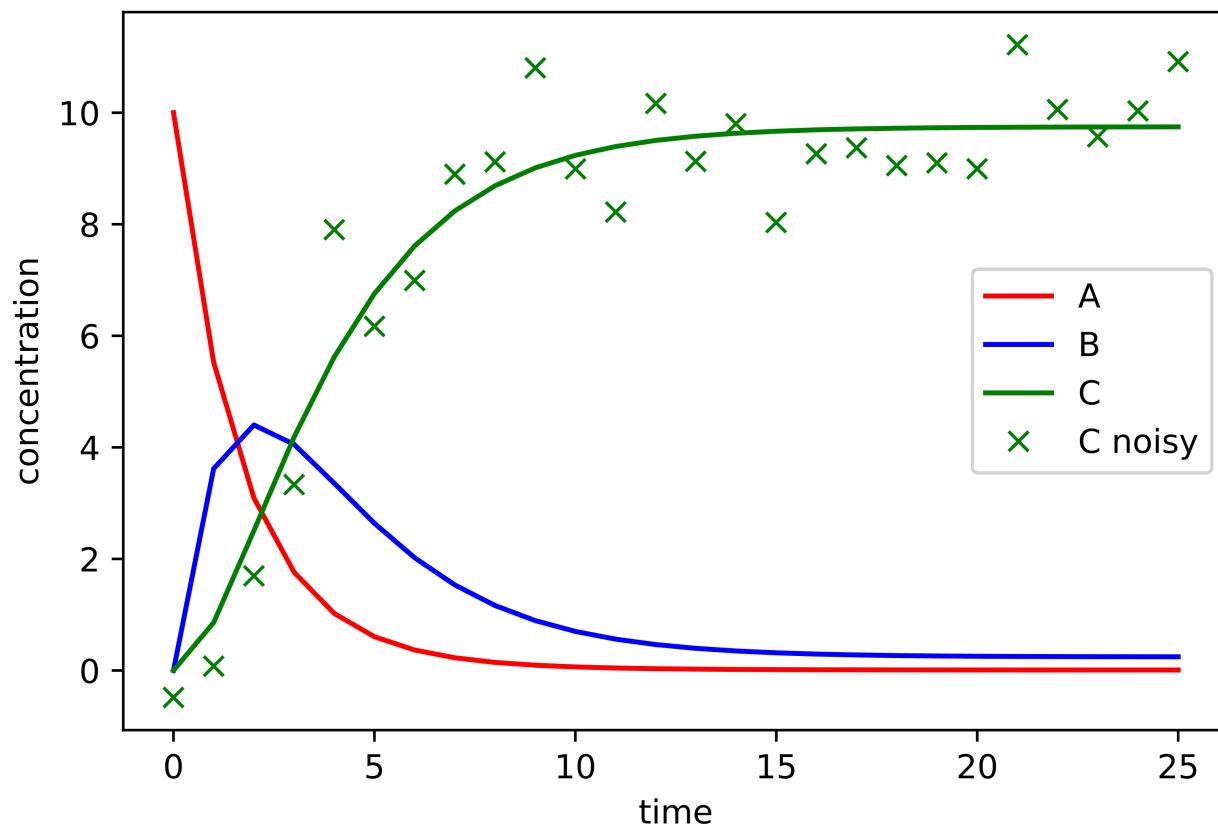
Simulate 'synthetic' data assuming

$$k_1 = 0.6, k_2 = 0.4, k_{-1} = 0.02, k_{-2} = 0.01$$

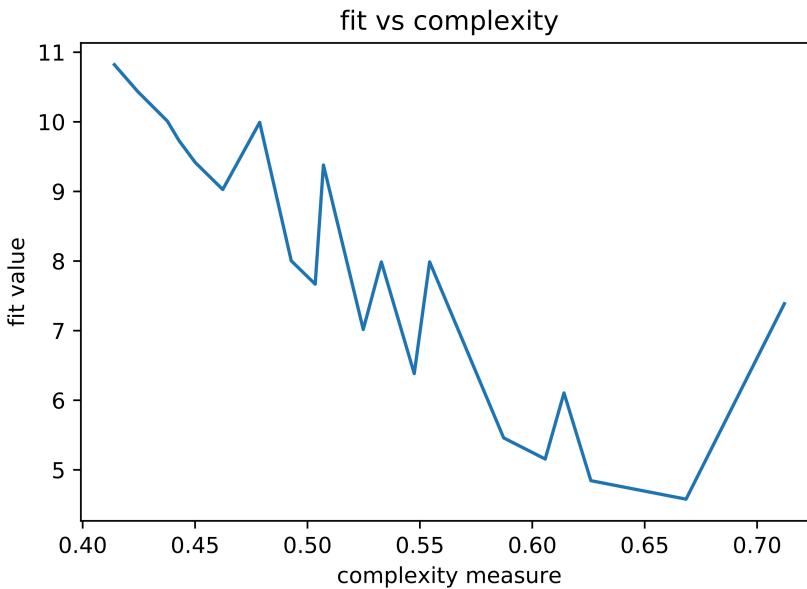
- Add Gaussian noise
- Assume can only measure $[C]$ (others unknown)

*Q: can we **recover** the original parameter values?*

SIMPLE REACTION: SYNTHETIC DATA

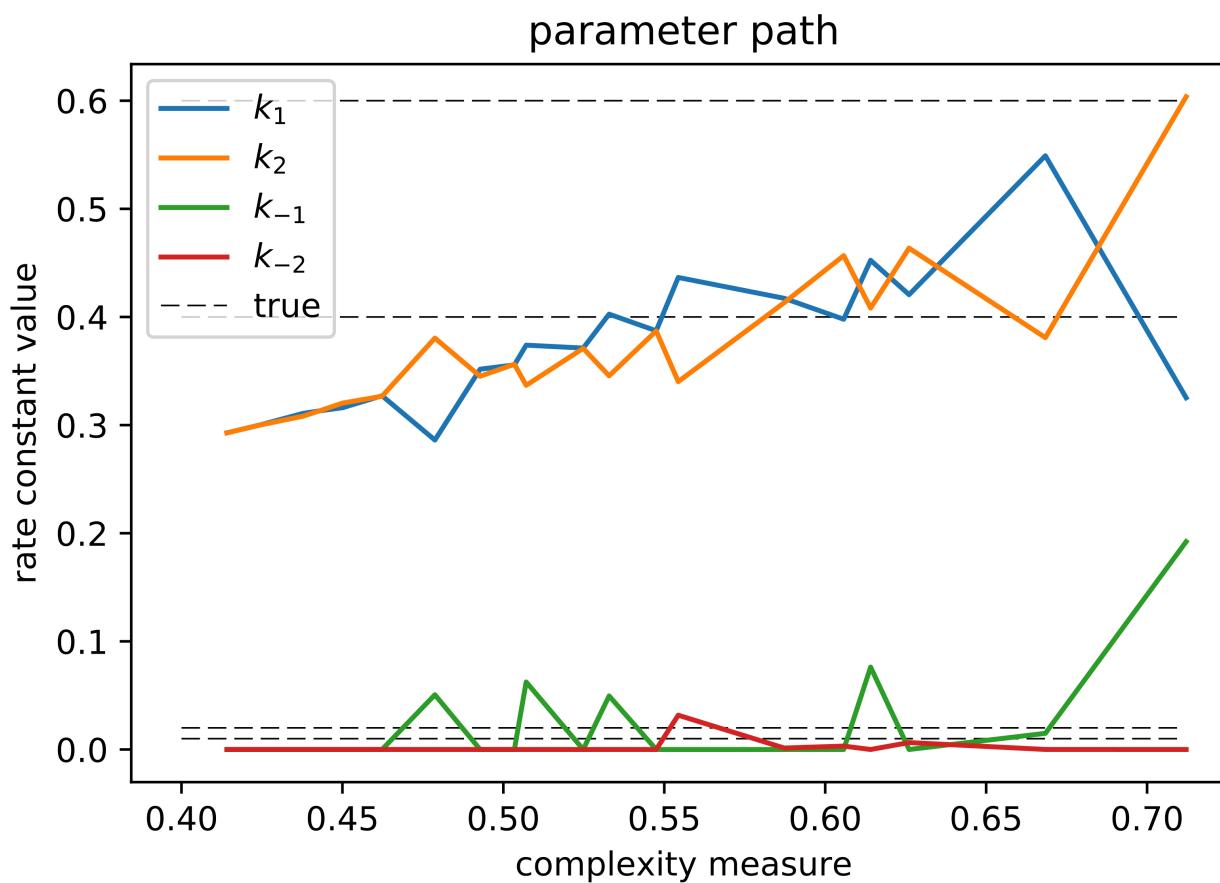


SIMPLE REACTION: FIT VS COMPLEXITY CURVE

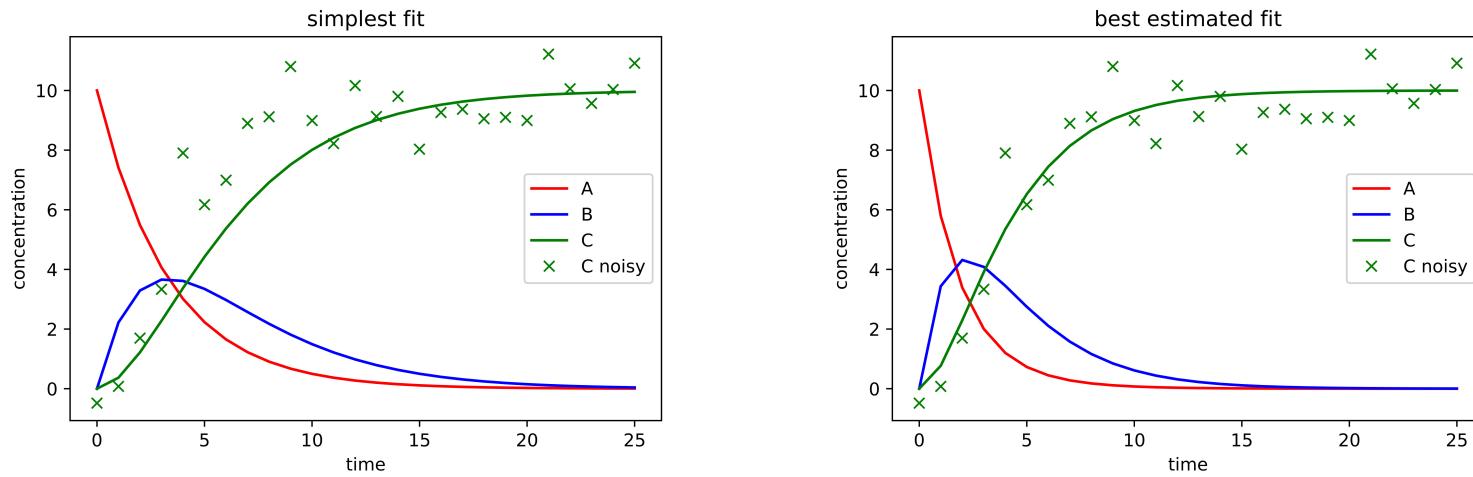


Note: not monotonically decreasing since get stuck in *local* minima in general.

SIMPLE REACTION: PARAMETER PATH



SIMPLE REACTION: BEST FIT AND SIMPLEST FIT



Simplest: $k_1 = 0.3, k_2 = 0.3, k_{-1} = 0.0, k_{-2} = 0.00$

Best: $k_1 = 0.55, k_2 = 0.38, k_{-1} = 0.015, k_{-2} = 0.00$

True: $k_1 = 0.6, k_2 = 0.4, k_{-1} = 0.02, k_{-2} = 0.01$

COMMENTS

- Can *recover* parameters from synthetic data OK
 - Some better than others
 - *Real* data much noiser and 'true' model unknown/non-existent
- Parameters can *trade-off* against each other
 - E.g. maybe only combinations are identifiable
- Need *efficient algorithms* for solving minimisation problems!

In general, parameter estimation requires care and awareness of trade-offs involved