

Intro to Inverse Problems

Andrea Arnold

Department of Mathematical Sciences
Worcester Polytechnic Institute



Mathematics Institute for Secondary Teaching
July 18, 2018

Motivational Example

Consider the line

$$y = f(x) = 3x + 2$$

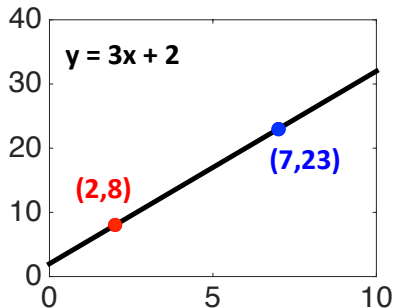
with slope $m = 3$ and y -intercept $b = 2$.

For any given x , we can compute the corresponding value of y !

$$x = 2 \Rightarrow y = f(2) = 8$$

$$x = 7 \Rightarrow y = f(7) = 23$$

$\vdots$



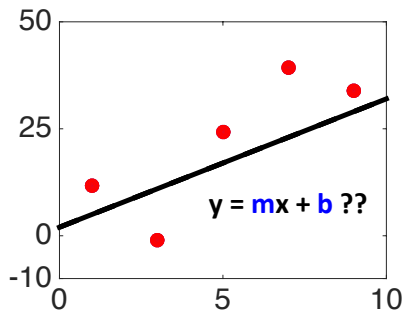
Motivational Example

Consider instead the points

$$(1, 11.72), \quad (3, -1.07), \quad (5, 24.17), \quad (7, 39.30), \quad (9, 33.89)$$

How can we find the equation of the line $y = mx + b$ that *best fits* the given data?

Which **parameters** m and b define the line of best fit?



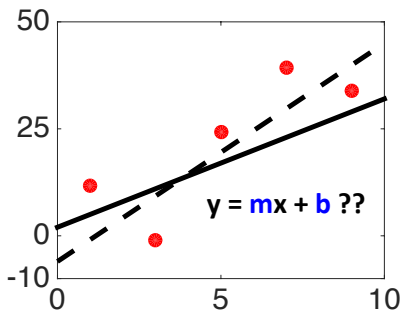
Motivational Example

Consider instead the points

$$(1, 11.72), \quad (3, -1.07), \quad (5, 24.17), \quad (7, 39.30), \quad (9, 33.89)$$

How can we find the equation of the line $y = mx + b$ that *best fits* the given data?

Which **parameters** m and b define the line of best fit?



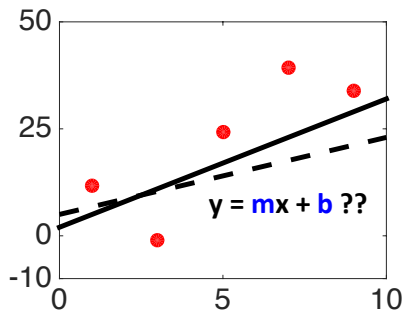
Motivational Example

Consider instead the points

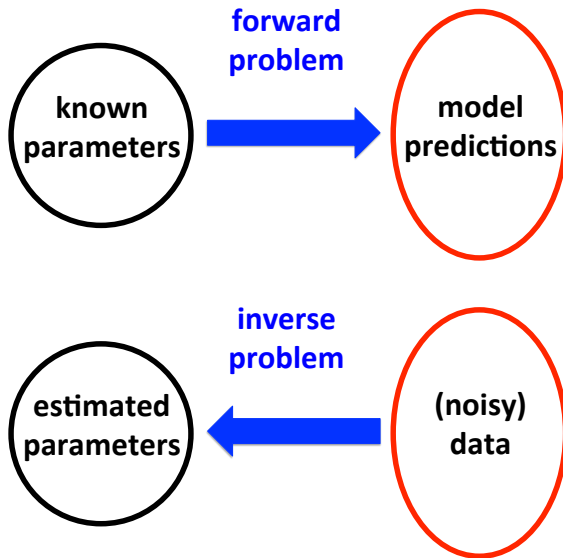
$$(1, 11.72), \quad (3, -1.07), \quad (5, 24.17), \quad (7, 39.30), \quad (9, 33.89)$$

How can we find the equation of the line $y = mx + b$ that *best fits* the given data?

Which **parameters** m and b define the line of best fit?



Forward vs. Inverse Problem



Forward vs. Inverse Problem

Given a map f explicitly relating two variables θ and y :

$$y = f(\theta)$$

- **Forward problem:** Compute y from known θ
- **Inverse problem:** Estimate unknown θ given y (or function of y) possibly corrupted by noise

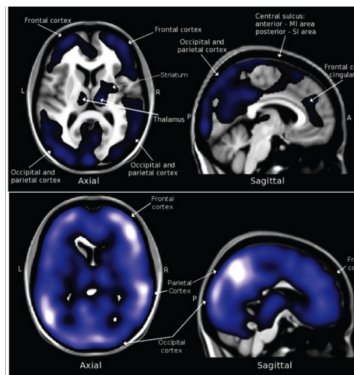
Model prediction = Forward problem

Parameter estimation = Inverse problem

Application Areas

Inverse problems arise in many fields of science and engineering

- Geophysics
- Signal processing
- Computer vision
- Machine learning
- Statistical inference
- Personalized medicine
- Medical imaging
- **Mathematical biology**

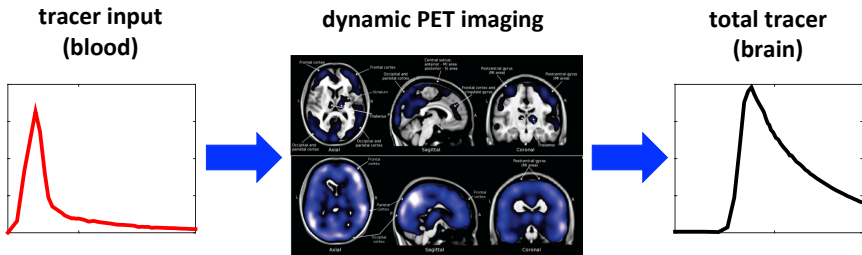


[1-¹¹C]-acetate PET imaging

Biological Example: Acetate Metabolism in Brain

Study in vivo neurodynamics in liver disease patients using model of neurochemical kinetics and $[1-^{11}\text{C}]$ -acetate PET measurements

- Does elevated level of ammonium in plasma primarily affect neurons or astrocytes?
- Acetate metabolized by astrocytes, not neurons!

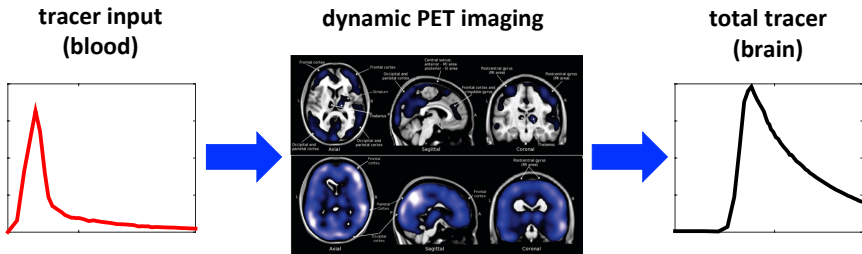


Iversen *et al.* (2009) *Gastroenterology*, 136, 863-871.

Biological Example: Acetate Metabolism in Brain

Data: 18 volunteers from 3 different groups of varying liver disease

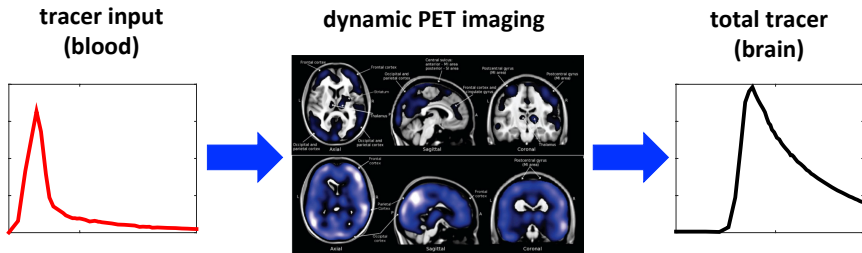
- 5 healthy control (HC)
- 7 cirrhotic liver (CL)
- 6 hepatic encephalopathy (HE)



Iversen *et al.* (2009) *Gastroenterology*, 136, 863-871.

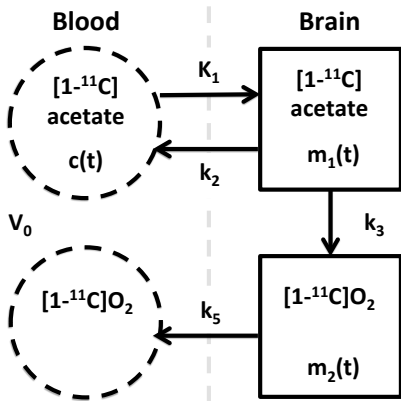
Biological Example: Acetate Metabolism in Brain

Question: Does the liver condition of the CL and HE patients significantly affect the astrocytic metabolism at the level of acetate uptake and clearance?



Iversen *et al.* (2009) *Gastroenterology*, 136, 863-871.

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (k_2 + k_3)m_1(t)$$

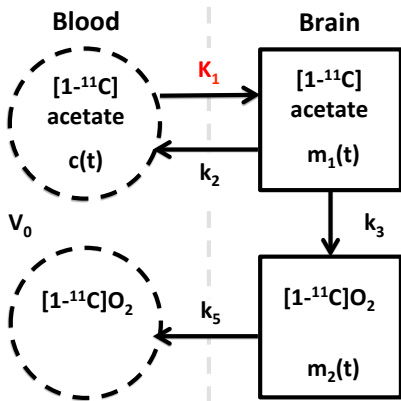
$$\frac{dm_2}{dt}(t) = k_3 m_1(t) - k_5 m_2(t)$$

Observation model:

$$m(t_j) = V_0 c(t_j) + m_1(t_j) + m_2(t_j)$$

Compartment model describing mass balance in the brain tissue

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = \textcolor{red}{K_1}c(t) - (k_2 + k_3)m_1(t)$$

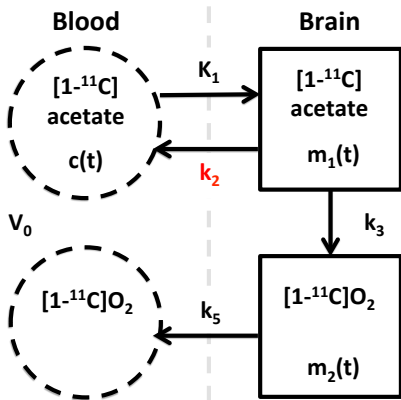
$$\frac{dm_2}{dt}(t) = k_3m_1(t) - k_5m_2(t)$$

Observation model:

$$m(t_j) = V_0c(t_j) + m_1(t_j) + m_2(t_j)$$

$\textcolor{red}{K_1}$ = clearance of tracer from blood to tissue

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (\textcolor{red}{k}_2 + k_3)m_1(t)$$

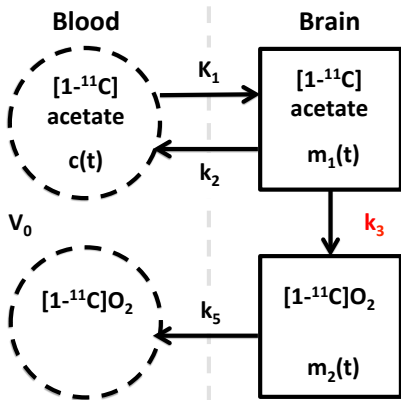
$$\frac{dm_2}{dt}(t) = k_3 m_1(t) - k_5 m_2(t)$$

Observation model:

$$m(t_j) = V_0 c(t_j) + m_1(t_j) + m_2(t_j)$$

$\textcolor{red}{k}_2, k_3, k_5 =$ transport rates between compartments

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (k_2 + \textcolor{red}{k}_3)m_1(t)$$

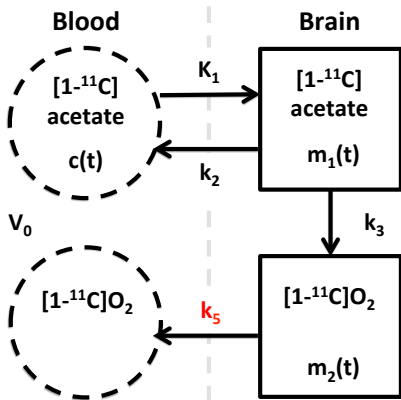
$$\frac{dm_2}{dt}(t) = \textcolor{red}{k}_3 m_1(t) - k_5 m_2(t)$$

Observation model:

$$m(t_j) = V_0 c(t_j) + m_1(t_j) + m_2(t_j)$$

$k_2, \textcolor{red}{k}_3, k_5$ = transport rates between compartments

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (k_2 + k_3)m_1(t)$$

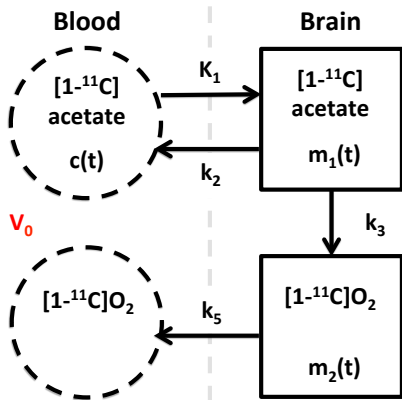
$$\frac{dm_2}{dt}(t) = k_3 m_1(t) - \textcolor{red}{k}_5 m_2(t)$$

Observation model:

$$m(t_j) = V_0 c(t_j) + m_1(t_j) + m_2(t_j)$$

$k_2, k_3, \textcolor{red}{k}_5$ = transport rates between compartments

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (k_2 + k_3) m_1(t)$$

$$\frac{dm_2}{dt}(t) = k_3 m_1(t) - k_5 m_2(t)$$

Observation model:

$$m(t_j) = \mathbf{V_0} c(t_j) + m_1(t_j) + m_2(t_j)$$

$\mathbf{V_0}$ = virtual volume of capillary

Example: Acetate Metabolism in Brain

Forward problem

- Given: Parameters K_1, k_2, k_3, k_5, V_0 + initial conditions
- Predict: $m_1(t)$ and $m_2(t)$ for $t > 0$

Inverse problem

- Given: Noisy measurements of total acetate concentration $m(t_i)$ at times $t_i, i = 1, \dots, n$
- Estimate: Parameters K_1, k_2, k_3, k_5, V_0

Solving Inverse Problems

- ① **Deterministic:** Solution is a *point estimate* of parameter values
 - Linear Least Squares
 - Nonlinear Least Squares
 - Weighted Least Squares
 - Regularization methods (e.g. Tikhonov)
- ② **Bayesian:** Solution is a *distribution* of possible parameter values
 - Non-sequential sampling (e.g. MCMC)
 - Sequential filtering

Least Squares Estimation

Find parameters θ so that model $f(x_i, \theta)$ “best fits” the given data y_i

$$y_i = f(x_i, \theta) + e_i, \quad i = 1, \dots, n$$

Solution minimizes the sum of the squares of the errors (residuals) between the model and observed data

$$e_i = y_i - f(x_i, \theta), \quad i = 1, \dots, n$$

Optimization!

Linear Least Squares

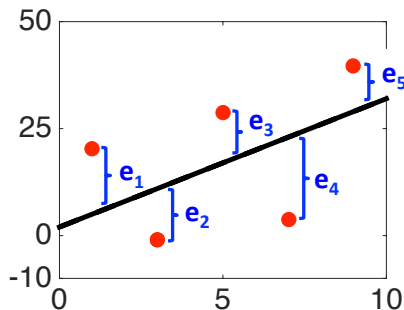
Example: Linear regression

Find the line of best fit $\hat{y} = mx + b$ to the given data!

- Given: Points (x_i, y_i) , $i = 1, \dots, n$
- Estimate: Parameters m and b that define line of best fit

Least squares approach:

Minimize the sum
of squared errors e_i !



Linear Least Squares

Inverse problem: Given the data (x_i, y_i) , $i = 1, \dots, n$, find the values of m and b that minimize the sum of the squared errors:

$$\min \sum_{i=1}^n e_i^2 \quad \equiv \quad \min S(m, b)$$

where

$$\begin{aligned} S(m, b) &:= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= \sum_{i=1}^n (y_i - mx_i - b)^2 \end{aligned}$$

Find m and b such that $\frac{\partial S}{\partial m} = 0$ and $\frac{\partial S}{\partial b} = 0$!

Linear Least Squares

2 equations with 2 unknowns, explicitly solve for m and b :

$$\left(\sum_{i=1}^n x_i^2\right)\mathbf{m} + \left(\sum_{i=1}^n x_i\right)\mathbf{b} = \sum_{i=1}^n x_i y_i$$

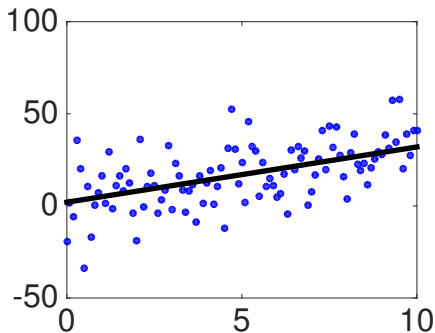
$$\left(\sum_{i=1}^n x_i\right)\mathbf{m} + n\mathbf{b} = \sum_{i=1}^n y_i$$

Least squares estimators:

$$\hat{\mathbf{m}} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}, \quad \hat{\mathbf{b}} = \frac{1}{n} \left(\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i \right)$$

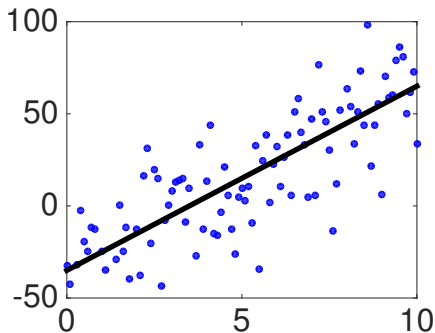
Linear Least Squares

Example: Linear regression



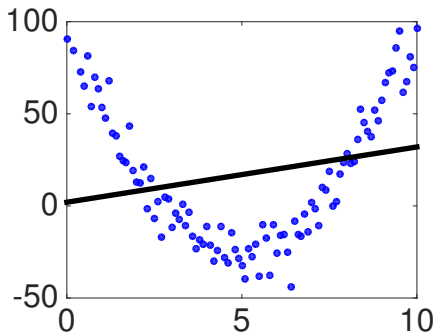
Linear Least Squares

Example: Linear regression



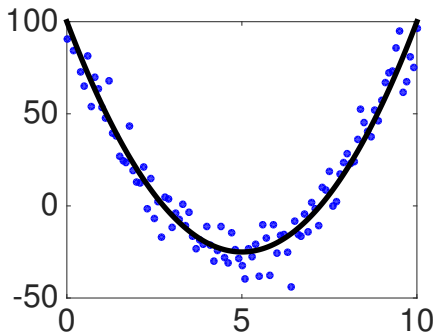
Linear Least Squares

Example: Linear regression



Linear Least Squares

Example: Quadratic regression



Linear Least Squares

General set-up:

$$\begin{aligned}y_i &= \theta_1 x_{i,1} + \theta_2 x_{i,2} + \cdots + \theta_p x_{i,p} + e_i, \quad i = 1, 2, \dots, n \\ &= \mathbf{x}_i^\top \boldsymbol{\theta} + e_i\end{aligned}$$

where

$$\mathbf{x}_i = \begin{bmatrix} x_{i,1} \\ \vdots \\ x_{i,p} \end{bmatrix} \in \mathbb{R}^p \quad \text{and} \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix} \in \mathbb{R}^p$$

Note: “Linear” refers to parameters, not independent variables!

Linear Least Squares

Matrix-vector form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{e}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n, \quad \mathbf{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} \in \mathbb{R}^n,$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^\top \\ \vdots \\ \mathbf{x}_n^\top \end{bmatrix} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,p} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,p} \end{bmatrix} \in \mathbb{R}^{n \times p}$$

Linear Least Squares

Inverse problem: Given the data $\mathbf{y}$, find

$$\hat{\theta} = \arg \min_{\theta} S(\theta)$$

where

$$S(\theta) = \sum_{i=1}^n (y_i - \mathbf{x}_i^T \theta)^2 = \|\mathbf{y} - \mathbf{X}\theta\|_2^2$$

Normal equations:

$$(\mathbf{X}^T \mathbf{X})\theta = \mathbf{X}^T \mathbf{y}$$

Least squares estimate:

$$\hat{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

Nonlinear Least Squares

Modification for nonlinearities in parameters:

$$y_i = f(\mathbf{x}_i, \theta) + e_i, \quad i = 1, \dots, n$$

where $f(\mathbf{x}_i, \theta)$ has nonlinear terms in θ

- No closed-form solution
- Must use iterative methods to estimate solution!

NLLS estimate:

$$\hat{\theta} = \arg \min_{\theta} S(\theta)$$

where

$$S(\theta) = \sum_{i=1}^n (y_i - f(\mathbf{x}_i, \theta))^2$$

Nonlinear Least Squares

Gradient:

$$\frac{\partial S}{\partial \theta_j} = 2 \sum_{i=1}^n e_i \frac{\partial e_i}{\partial \theta_j}, \quad j = 1, \dots, p$$

Use iterative methods to approximate $\hat{\theta}$ that satisfies $\frac{\partial S}{\partial \theta_j} = 0$!

Start with initial guess θ^0 , then update: At step $k + 1$, the new estimate is

$$\theta^{k+1} = \theta^k + \Delta \theta$$

where the previous estimate θ^k is corrected by $\Delta \theta$

Nonlinear Least Squares

Jacobian:

$$\mathbf{J} = \left[\frac{\partial e_i}{\partial \theta_j} \right] \in \mathbb{R}^{n \times p}, \quad i = 1, \dots, n, \quad j = 1, \dots, p$$

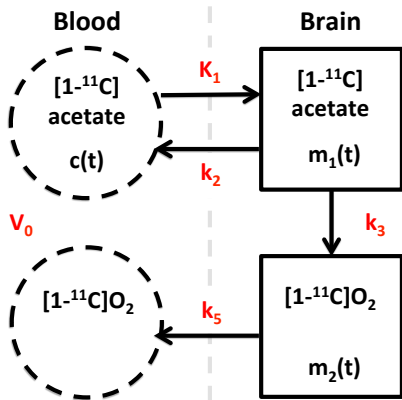
Normal equations for NLLS:

$$(\mathbf{J}^T \mathbf{J}) \Delta \theta = \mathbf{J}^T \mathbf{e}$$

Gauss-Newton step:

$$\Delta \theta = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \mathbf{e}$$

Example: Acetate Metabolism in Brain



Forward model:

$$\frac{dm_1}{dt}(t) = K_1 c(t) - (k_2 + k_3) m_1(t)$$

$$\frac{dm_2}{dt}(t) = k_3 m_1(t) - k_5 m_2(t)$$

Observation model:

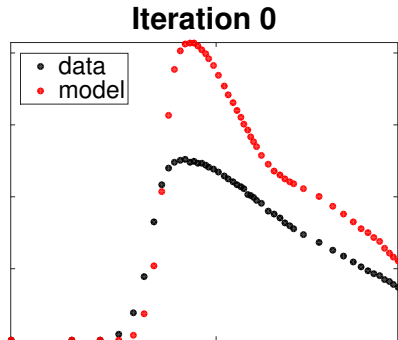
$$m(t_j) = V_0 c(t_j) + m_1(t_j) + m_2(t_j)$$

Estimate K_1, k_2, k_3, k_5, V_0 !

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 8.0000e-04 \\K_1 &= 9.6000e-03 \\k_2 &= 1.6000e-02 \\k_3 &= 1.6000e-03 \\k_5 &= 1.6000e-03\end{aligned}$$

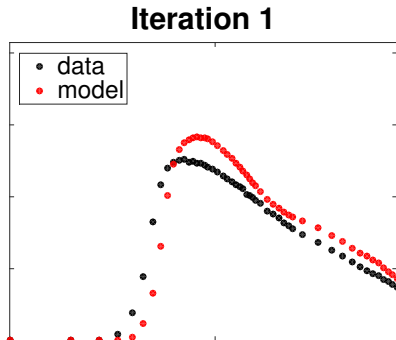


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 1.0000e-04 \\K_1 &= 6.0301e-03 \\k_2 &= 1.2315e-02 \\k_3 &= 1.5098e-03 \\k_5 &= 1.5859e-03\end{aligned}$$

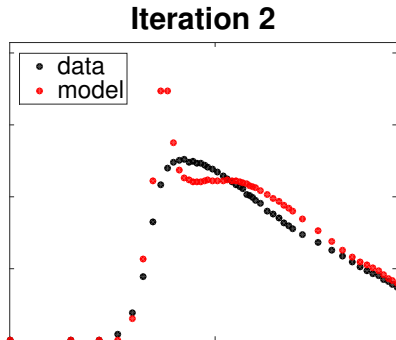


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 9.8783e-02 \\K_1 &= 2.9220e-03 \\k_2 &= 5.1900e-03 \\k_3 &= 3.7134e-04 \\k_5 &= 1.1363e-03\end{aligned}$$

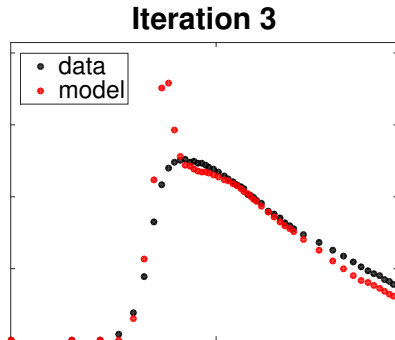


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 9.6125e-02 \\K_1 &= 3.5245e-03 \\k_2 &= 8.5702e-03 \\k_3 &= 1.9830e-03 \\k_5 &= 3.5272e-03\end{aligned}$$

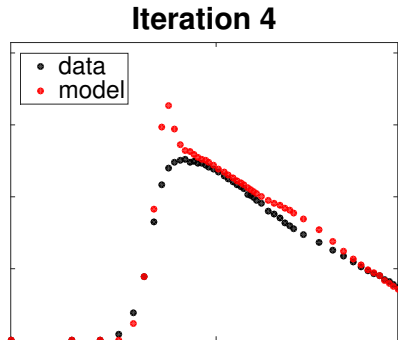


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 6.8416e-02 \\K_1 &= 4.7904e-03 \\k_2 &= 1.4403e-02 \\k_3 &= 6.3611e-03 \\k_5 &= 3.2959e-03\end{aligned}$$

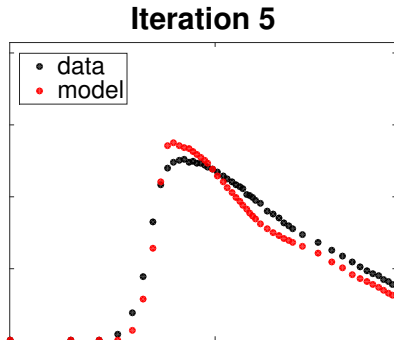


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 3.5012e-02 \\K_1 &= 5.7135e-03 \\k_2 &= 1.6762e-02 \\k_3 &= 2.3035e-03 \\k_5 &= 2.4845e-03\end{aligned}$$

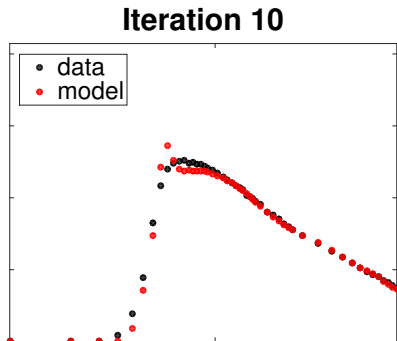


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 5.3690e-02 \\K_1 &= 4.1074e-03 \\k_2 &= 9.6078e-03 \\k_3 &= 1.4268e-03 \\k_5 &= 1.7815e-03\end{aligned}$$

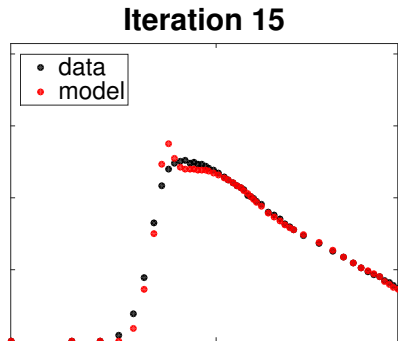


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 5.4292e-02 \\K_1 &= 4.1699e-03 \\k_2 &= 9.8703e-03 \\k_3 &= 1.5060e-03 \\k_5 &= 1.8212e-03\end{aligned}$$

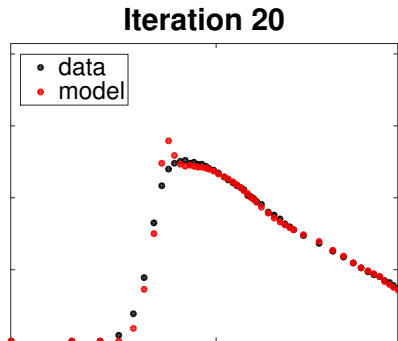


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 5.3836e-02 \\K_1 &= 4.3073e-03 \\k_2 &= 1.0373e-02 \\k_3 &= 1.6380e-03 \\k_5 &= 1.8854e-03\end{aligned}$$

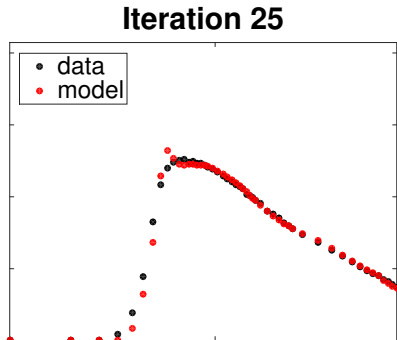


Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain

Use NLLS to find parameter values that best fit data

$$\begin{aligned}V_0 &= 4.6091e-02 \\K_1 &= 4.4525e-03 \\k_2 &= 1.0555e-02 \\k_3 &= 1.6081e-03 \\k_5 &= 1.8802e-03\end{aligned}$$



Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Deterministic vs. Bayesian Solutions

Deterministic methods (like least squares) result in a single point estimate of parameter values

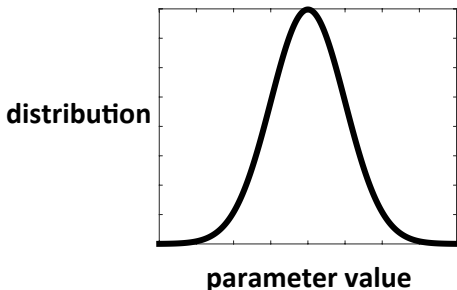
- No automatic measure of uncertainty in solution or corresponding model predictions
- Need to apply, e.g., frequentist statistics formulas

Bayesian methods (like filtering) result in a distribution of possible parameter values

- Provides natural framework for including uncertainties
- Prior information can be easily encoded
- Automatic measure of reliability in solution

Bayesian Solution

Bayesian approach to solving inverse problems involves modeling all unknown quantities as **random variables**



- Randomness described in terms of probability distributions
- Distributions convey what is known about parameter values and to what extent
- Provides a natural link between inverse problems and statistical inference

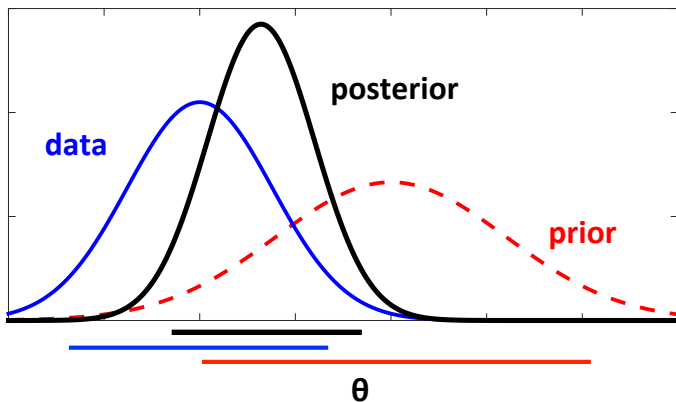
Bayes' theorem:

$$\pi(\theta \mid \mathbf{y}) \propto \pi(\mathbf{y} \mid \theta)\pi(\theta)$$

where

- $\pi(\theta \mid \mathbf{y})$ is the posterior density (solution) of parameters θ given the data $\mathbf{y}$
- $\pi(\mathbf{y} \mid \theta)$ is the likelihood of obtaining the data $\mathbf{y}$ given the parameters θ
- $\pi(\theta)$ is the prior density, encoding any known information about the parameters θ

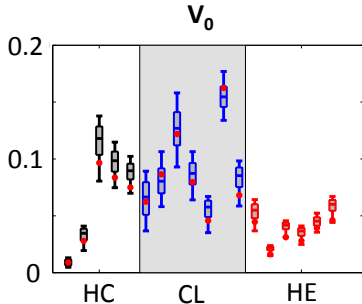
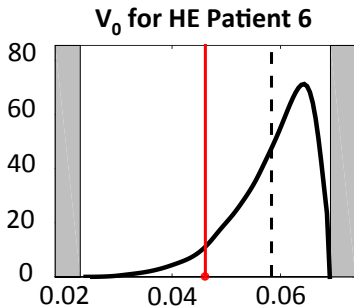
Bayesian Solution



Example: Acetate Metabolism in Brain

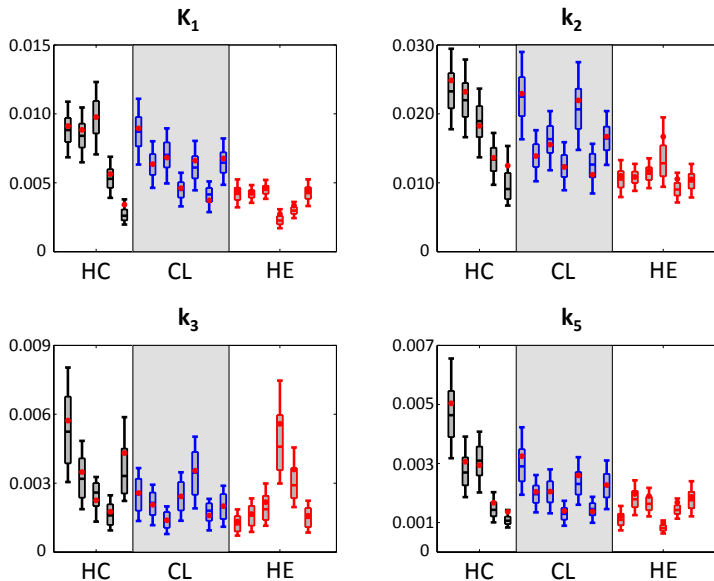
Apply nonlinear Bayesian filter to estimate parameter distributions

- Use NLLS parameter values to inform prior!



Arnold *et al.* (2015) *Mathematical Medicine & Biology*, 32 (4), 367-382

Example: Acetate Metabolism in Brain



Conclusions

Deterministic and Bayesian techniques for solving inverse problems

- Application to mathematical biology
- Not mutually exclusive – combination of techniques can be useful!

References

Inverse problems texts (deterministic)

- P. C. Hansen (1998) *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Mathematical Modeling and Computation. SIAM, Philadelphia.
- H. W. Engl, M. Hanke and A. Neubauer (2000) *Regularization of Inverse Problems*. Mathematics and Its Applications. Kluwer Academic Publishers, Boston.
- H. T. Banks, S. Hu and W. C. Thompson (2014) *Modeling and Inverse Problems in the Presence of Uncertainty*. CRC Press, New York.

References

Inverse problems texts (Bayesian)

- J. P. Kaipio and E. Somersalo (2005) *Statistical and Computational Inverse Problems*. Applied Mathematical Sciences. Springer, New York.
- D. Calvetti and E. Somersalo (2007) *Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*. Surveys and Tutorials in the Applied Mathematical Sciences. Springer, New York.

Optimization texts

- J. E. Dennis Jr. and R. B. Schnabel (1996) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Classics in Applied Mathematics. SIAM, Philadelphia.
- C. T. Kelley (1999) *Iterative Methods for Optimization*. SIAM, Philadelphia.

Acetate metabolism application

- P. Iversen, M. Sorensen, L. K. Bak, H. S. Waagepetersen, M. S. Vafaei, P. Borghammer and S. Keiding (2009) Low cerebral oxygen consumption and low flow in patients with cirrhosis and an acute episode of hepatic encephalopathy. *Gastroenterology*, 136, 863-871.
- A. Gjedde, S. Keiding, H. Vilstrup and P. Iversen (2010) No oxygen delivery limitation in hepatic encephalopathy. *Metabolic Brain Disease*, 25, 57-63.
- A. Arnold, D. Calvetti, A. Gjedde, P. Iversen and E. Somersalo (2015) Astrocytic tracer dynamics estimated from $[1-^{11}\text{C}]$ -acetate PET measurements. *Mathematical Medicine & Biology*, 32 (4), 367-382.

Bayesian filtering

- A. Arnold, D. Calvetti and E. Somersalo (2013) Linear multistep methods, particle filtering and sequential Monte Carlo. *Inverse Problems*, 29 (8), 085007.
- A. Arnold, D. Calvetti and E. Somersalo (2014) Parameter estimation for stiff deterministic dynamical systems via ensemble Kalman filter. *Inverse Problems*, 30 (10), 105008.
- A. Arnold, D. Calvetti and E. Somersalo (2015) Vectorized and parallel particle filter SMC parameter estimation for stiff ODEs. In: *Dynamical Systems and Differential Equations, AIMS Proceedings 2015* (M. de León, W. Feng, Z. Feng, X. Lu, J. M. Martell, J. Parcet, D. Peralta-Salas and W. Ruan eds) Madrid, Spain, 75-84.

References

Bayesian filtering

- A. Arnold (2014) Sequential Monte Carlo parameter estimation for differential equations. *Ph.D. Thesis*, Case Western Reserve University, Cleveland, OH.
- A. Arnold, C. Battista, D. Bia, Y. Zocalo German, R. L. Armentano, H. Tran and M. S. Olufsen (2016) Uncertainty quantification in a patient-specific 1D arterial network model: EnKF-based inflow estimator. *ASME Journal of Verification, Validation and Uncertainty Quantification*, 2 (1), 011002.
- A. Arnold and A. L. Lloyd (2018) An approach to periodic, time-varying parameter estimation using nonlinear filtering. Accepted, to appear in *Inverse Problems*.

Thank you!