# Introduction to modeling, simulation and data science in oncology

S. Benzekry

Formation « Ecole doctorale Mathématiques et Informatique »

MONC

Modeling in ONCology

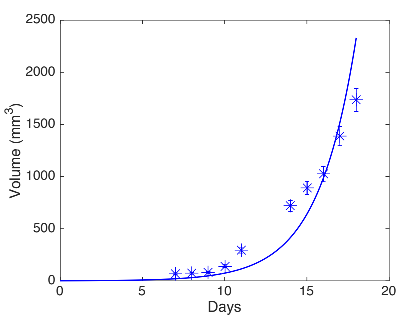
# Cancer: a major public health concern

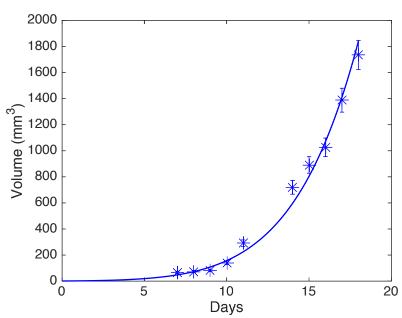
- Second leading **cause of death** worldwide (1 in 6 deaths, 8.8 million deaths in 2015)
- **First** cause of death in France (> 1 in 4 deaths) *InVS and INCa, 2011*
- Cumulative risks of **developing** a cancer: 30.9% in males and 23.3% in females
- Cumulative risks of **death** by cancer: 14.3% and 9%
- Most prevalent cancer types: **breast** in women, **prostate** in men
- Largest number of deaths: **lung** cancer
- One third of deaths from cancer are due to 5 leading **behavioral and dietary risks**: **tobacco** use (22%), high body mass index, low fruit and vegetable intake, lack of physical activity and alcohol use

# Can mathematical models be of help in oncology?

## Understand (biology)

- **Theoretical** framework for description of the process
- Test different **hypotheses** and reject non-valid ones




Exponential

$$\frac{dV}{dt} = aV$$

Power law

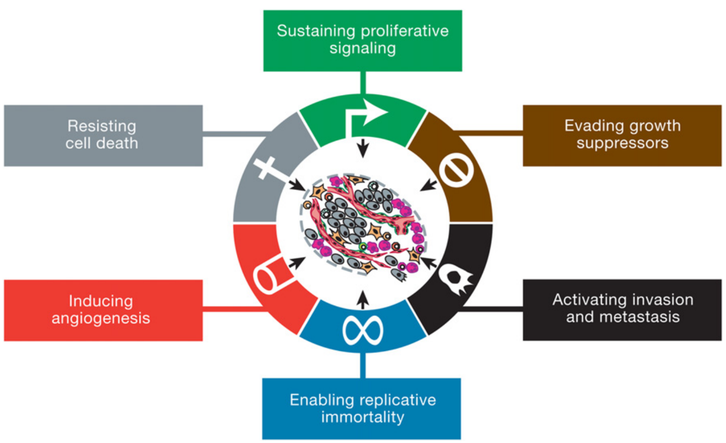
$$\frac{dV}{dt} = aV^{\gamma}$$

*Benzekry et al., PLoS Comp Biol, 2014*

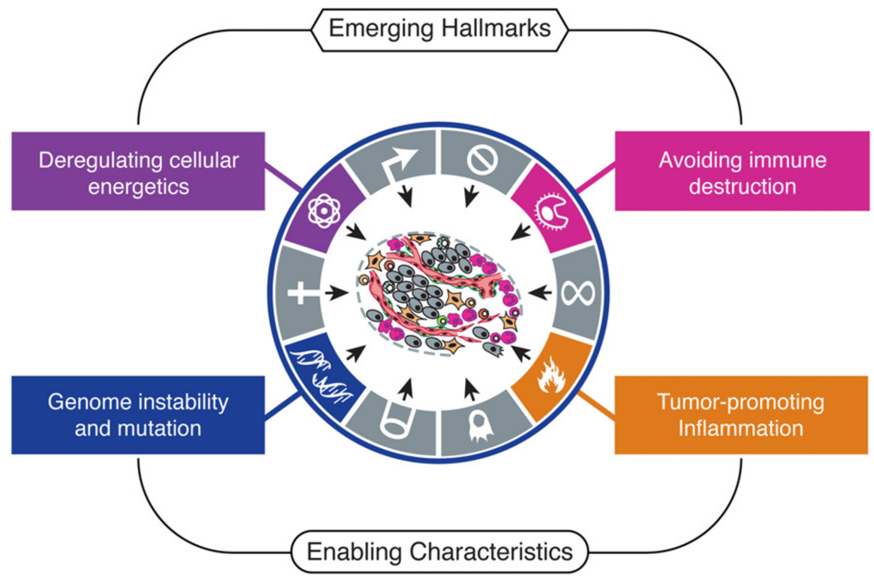
# What is a cancer?

- Tumor = malignant **neoplasm**. neo = new, plasma = formation
- Usually assumed that it departs from a cell undergoing several genetic and epigenetic changes leading to **abnormal proliferation**

## Hallmarks of cancer

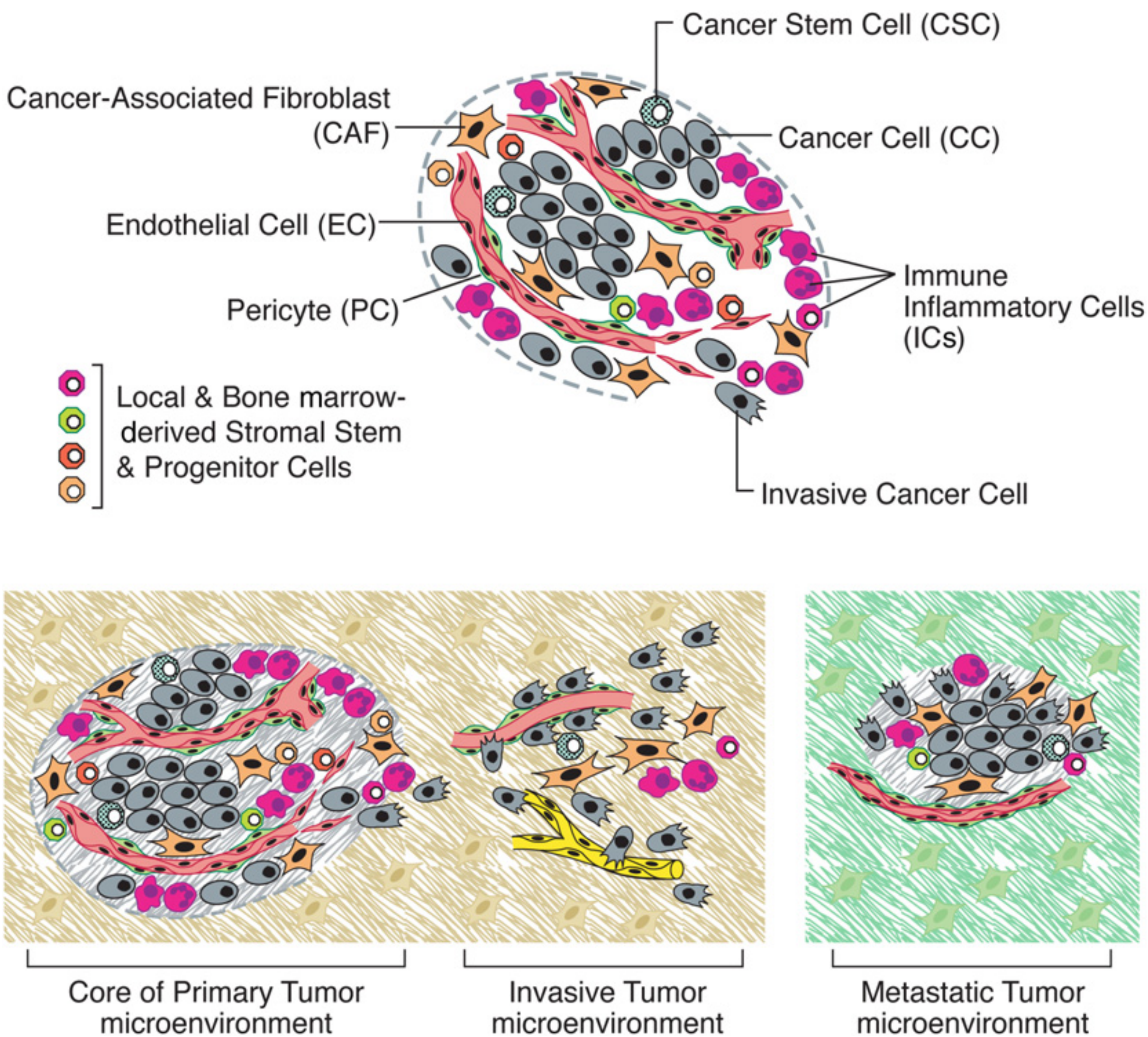



*Hanahan and Weinberg, Cell, 2000*



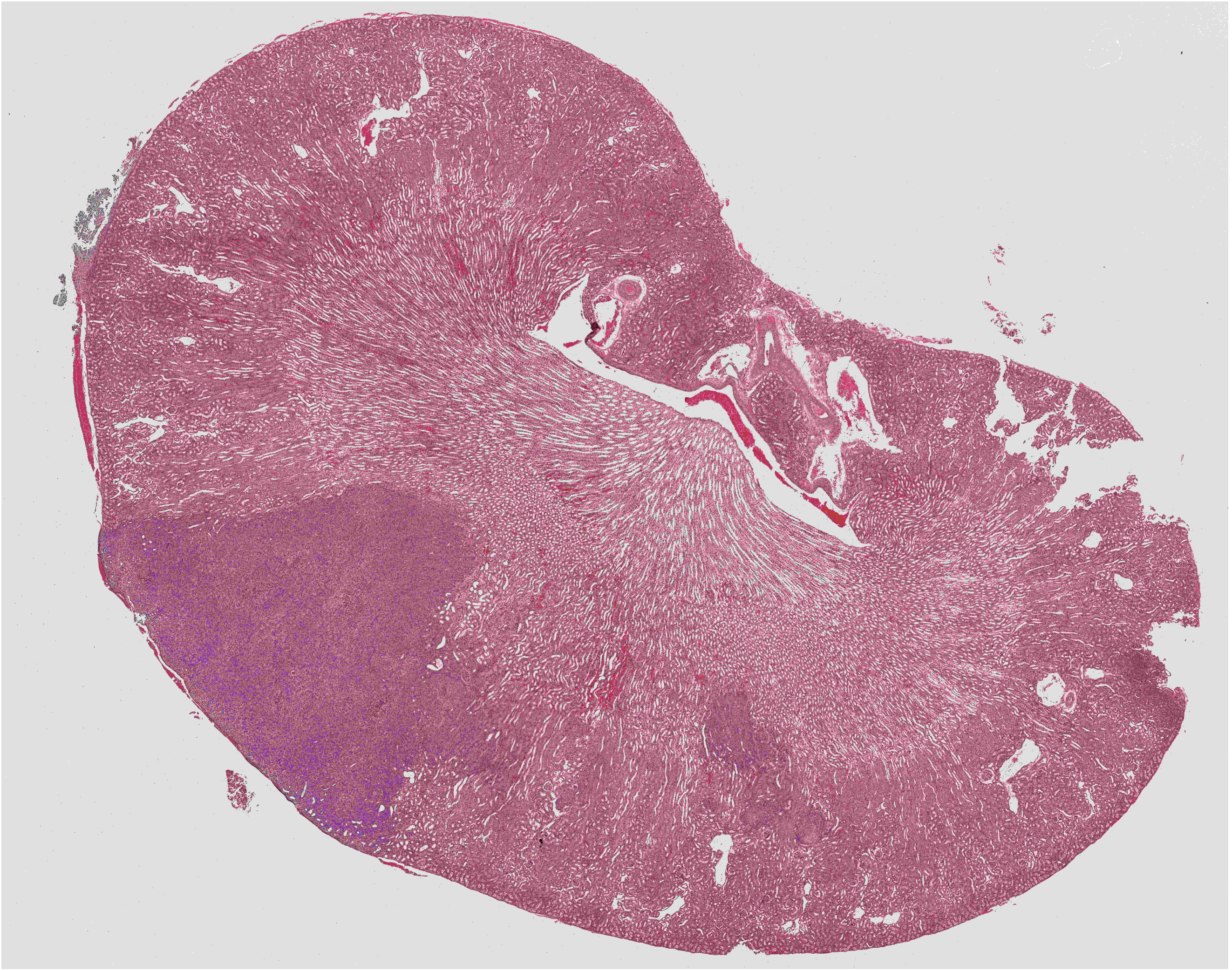

*Hanahan and Weinberg, Cell, 2011*

# Microenvironment




*Hanahan and Weinberg, Cell, 2011*

# A kidney tumor observed by Hematoxylin and Eosin staining

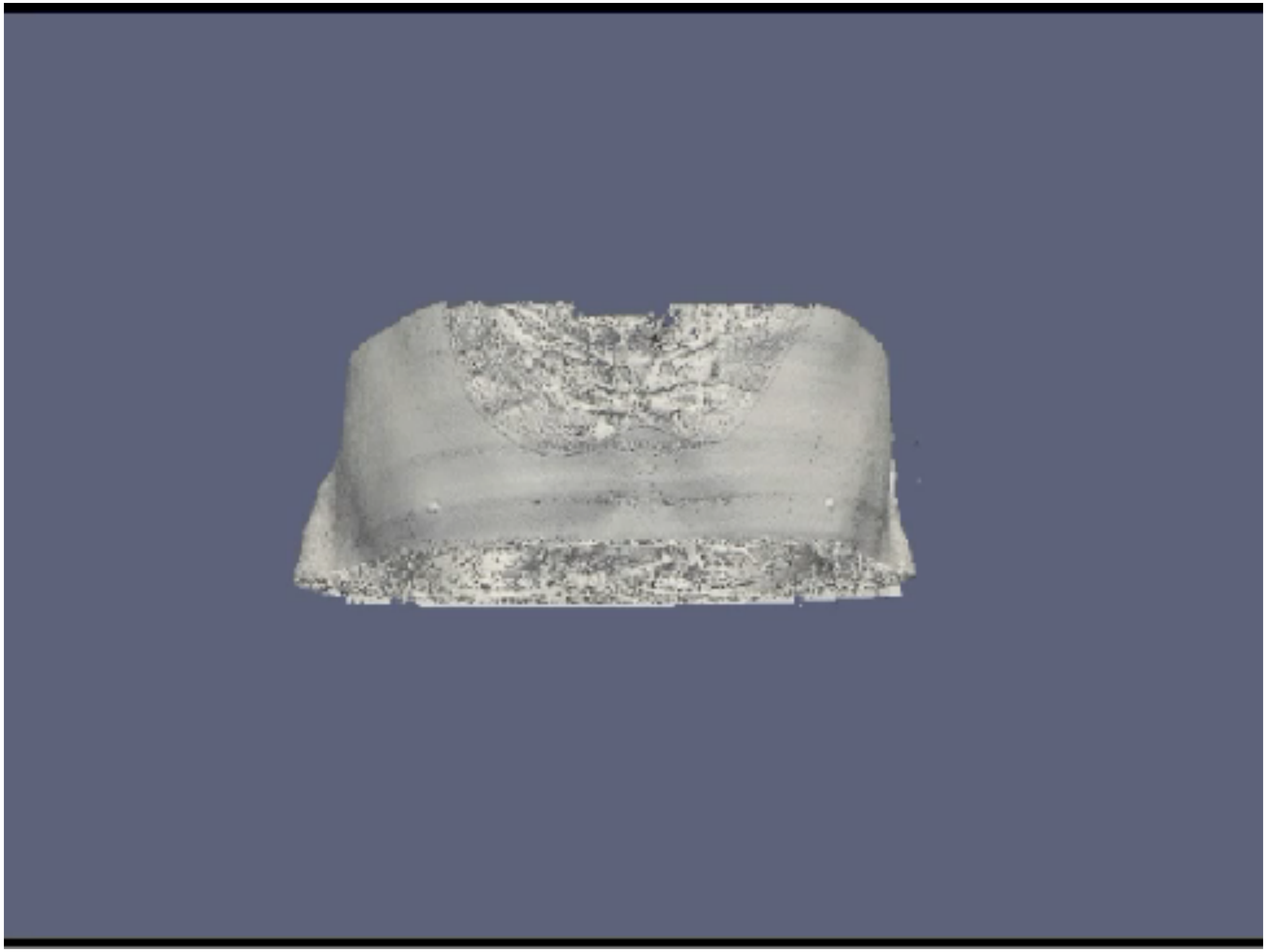


We will focus here on **carcinomas:** solid cancers from epithelial origin

# Can mathematical models be of help in oncology?

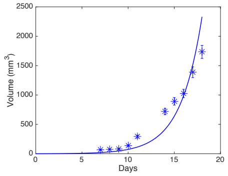
## Predict and control (clinic)

- Predict tumor growth



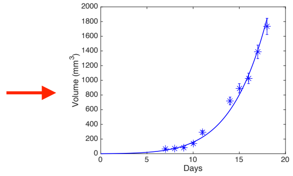
### Understand (biology)

- **Theoretical** framework for description of the process
- Test different **hypotheses** and reject non-valid ones



Exponential

$$\frac{dV}{dt} = aV$$



Power law

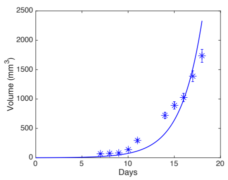
$$\frac{dV}{dt} = aV^{\gamma}$$

*Jouganous, Colin, Saut et al., 2014*

# Can mathematical models be of help in oncology?

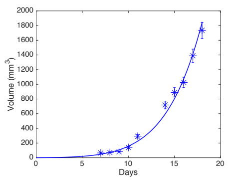
### Understand (biology)

- **Theoretical** framework for description of the process
- Test different **hypotheses** and reject non-valid ones




Exponential

$$\frac{dV}{dt} = aV$$

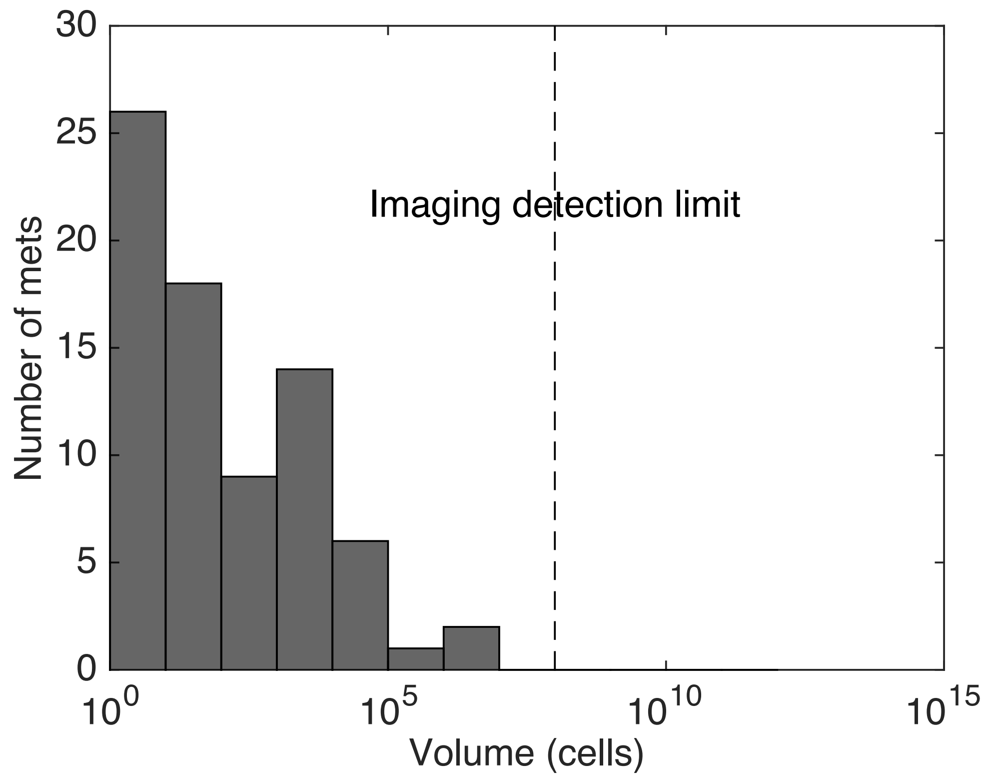



Power law

$$\frac{dV}{dt} = aV^{\gamma}$$

### Predict and control (clinic)

- Predict **metastasis**
- **Personalize** (adjuvant) therapy

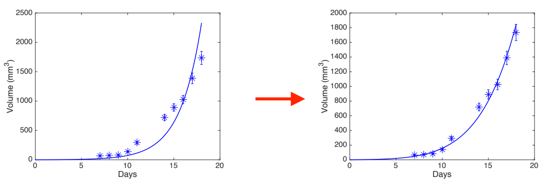



*Benzekry et al., Cancer Res, 2016*

# Can mathematical models be of help in oncology?

## Understand (biology)

- **Theoretical** framework for description of the process
- Test different **hypotheses** and reject non-valid ones




Exponential

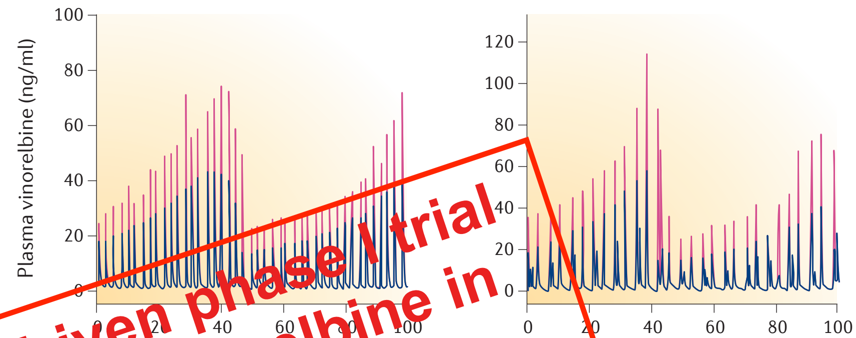
Power law

## Predict and control (clinic)

- Rational and individual design of **drug regimen**

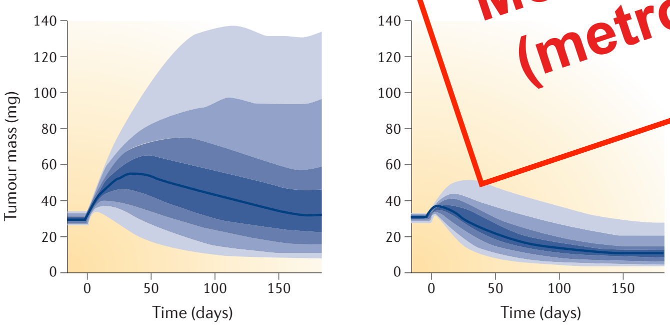
**Empirical dosing**
**D1-D3-D5 50 mg**

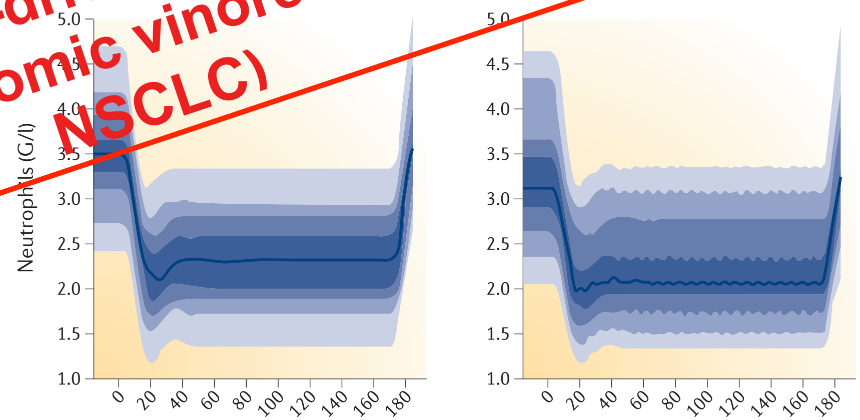
**Model-based dosing**
**D1-D2-D4 60-30-60 mg**




PK

Efficacy

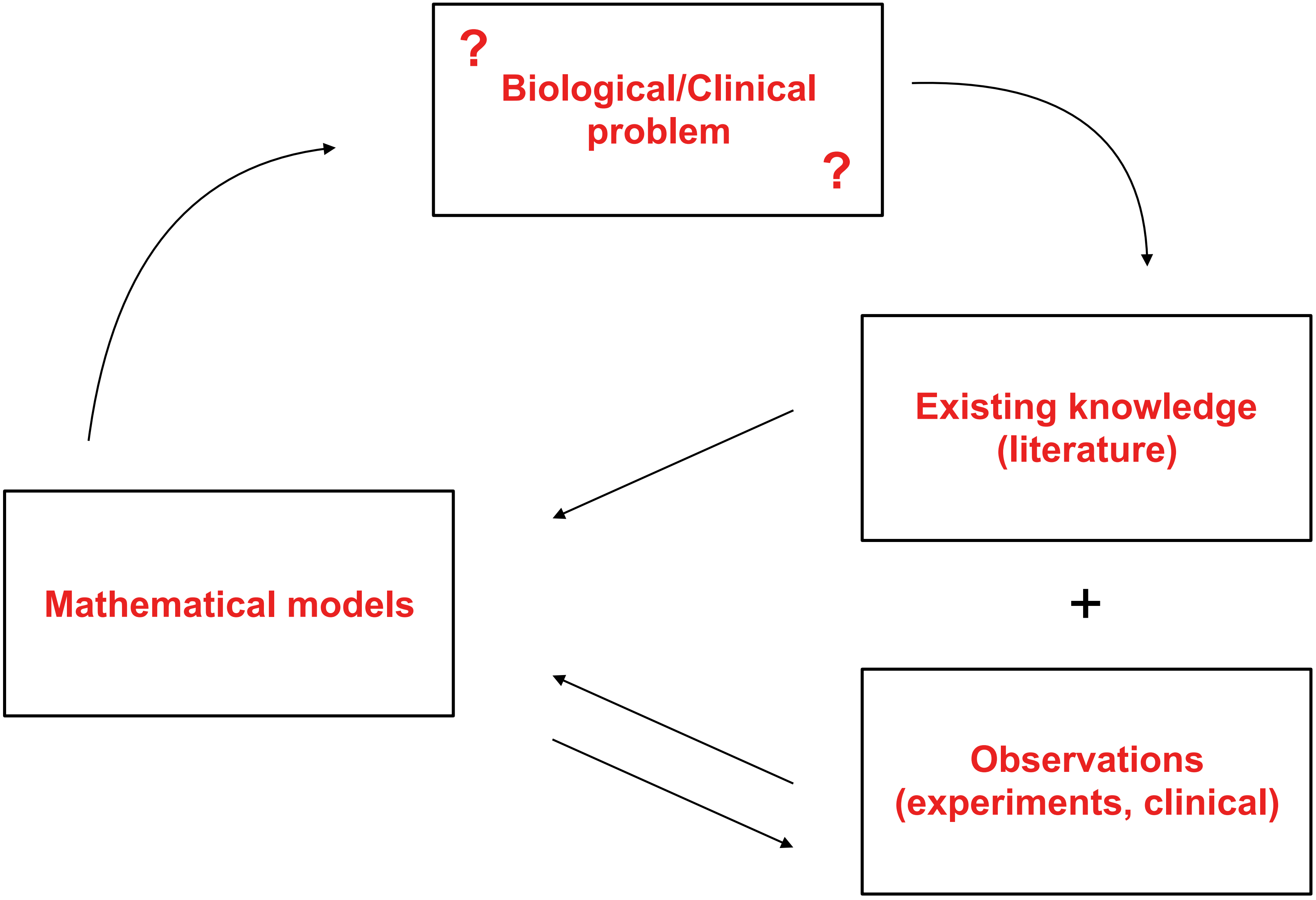



Toxicity

**Modeling-driven phase I trial (metronomic vinorelbine in NSCLC)**

*Barbolosi et al., Nat Rev Clin Onco., 2015*


?
Biological/Clinical
problem
?
Existing knowledge
(literature)
+
Mathematical models
Observations
(experiments, clinical)

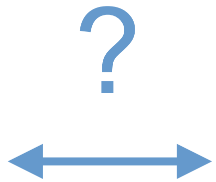
# 1. Fitting a model

# 1.1 Fitting a *linear* model

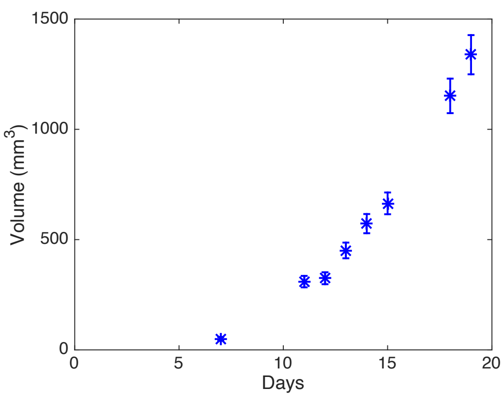
# Data

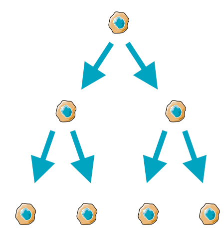
# Theory

# Mathematical model

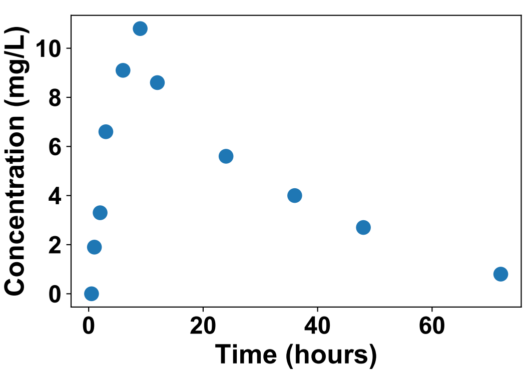


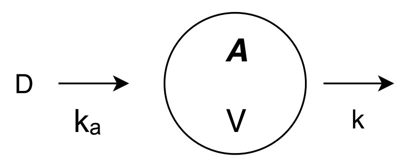

$$M : \begin{array}{ccc} \mathbb{R} \times \mathbb{R}^p & \to & \mathbb{R} \\ (t, \theta) & \mapsto & M(t, \theta) \end{array}$$



$$M(t, \theta) = e^{\theta t}$$



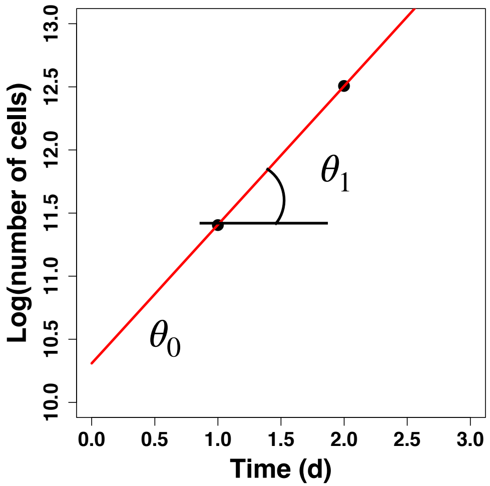




$$\begin{cases} \frac{dA_a}{dt} = -k_a A_a \\ \frac{dA}{dt} = k_a A_a - kA \\ A_a(t=0) = D, \quad (t=0) = 0 \end{cases}$$

$$C(t) = \frac{A(t)}{V}.$$

# Linear system: Equation of a line

$$y = \theta_0 + \theta_1 t$$




$$\begin{cases} y_1 = 1 \times \theta_0 + t_1 \times \theta_1 \\ y_2 = 1 \times \theta_0 + t_2 \times \theta_1 \end{cases} \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

$$y = M \cdot \theta \Rightarrow \theta = M^{-1} \cdot y$$

$$\begin{cases} 11.4 = 1 \times \theta_0 + 1 \times \theta_1 \\ 12.5 = 1 \times \theta_0 + 2 \times \theta_1 \end{cases} \Leftrightarrow \begin{pmatrix} 11.4 \\ 12.5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

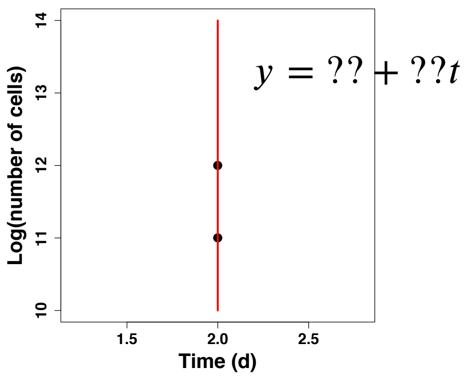
$$M^{-1} \text{??} \qquad M^{-1} := \text{“}\frac{1}{M}\text{”}, \quad M \cdot M^{-1} = \text{“}1\text{”} = I$$

$$\theta_0 = 10.3, \ \theta_1 = 1.1$$

is $M \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ sufficient?

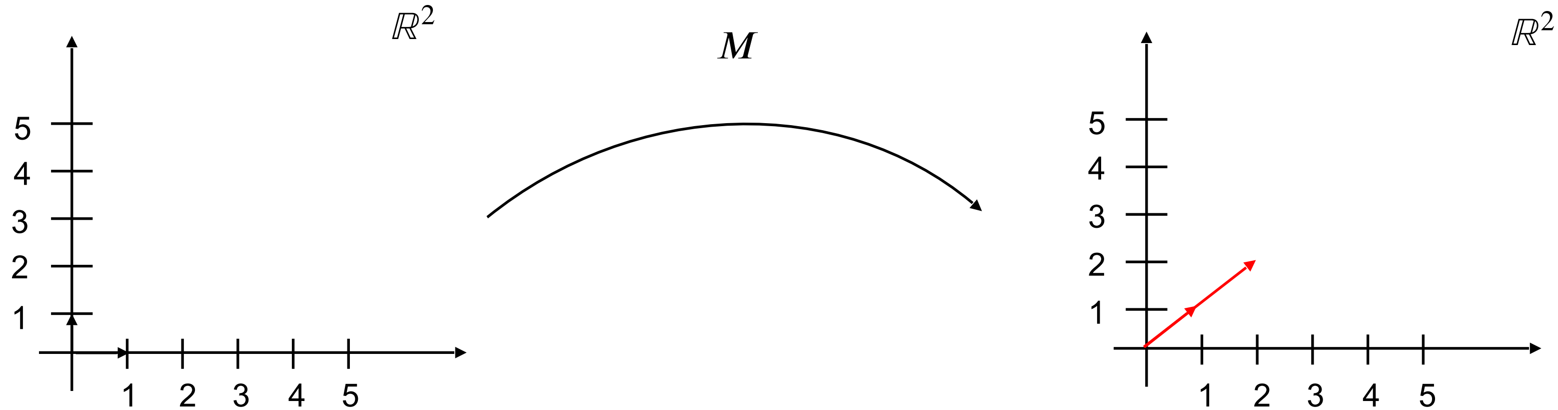
Doubling time = $\dfrac{\ln 2}{\theta_1} \times 24 = 15.1$ hours

# Invertible matrix

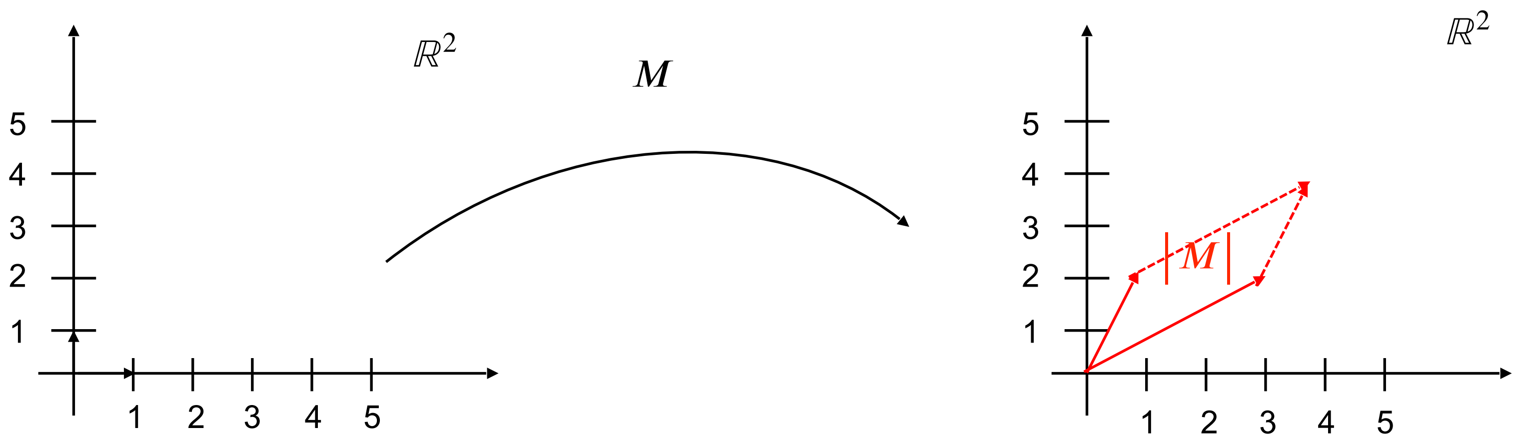



$$\begin{cases} 11 = \theta_0 + 2 \times \theta_1 \\ 12 = \theta_0 + 2 \times \theta_1 \end{cases} \Leftrightarrow \begin{pmatrix} 11 \\ 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

$M = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ is not invertible because its column (and row) vectors are colinear



# Determinant




- The determinant of $M$, denoted $\left| M \right|$, is the area of the parallelogramm spanned by the column vectors of $M$

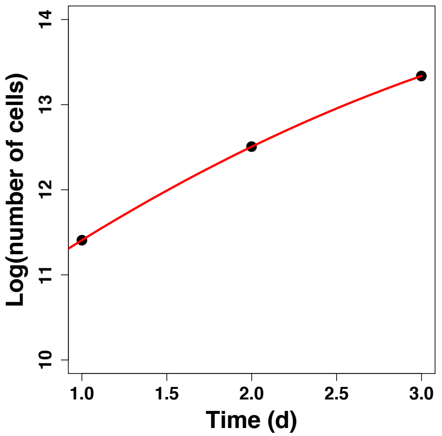
- For $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ it is given by $ad - bc$.

- It can be generalized in any dimension and is a measure of the colinearity (and correlation) of the vectors

- $\left| M \right| \neq 0 \Leftrightarrow$ $M$ is invertible $\Leftrightarrow$ the column (and row) vectors of $M$ are independent

# Linear system: polynomial interpolation

- What if we have 3 points?
- 3 points ⇔ 3 degrees of freedom ⇔ 3 parameters




$$y = \theta_0 + \theta_1 t + \theta_2 t^2$$

$$y = 10 + 1.5t - 0.13t^2$$

3 unknowns

3 equations

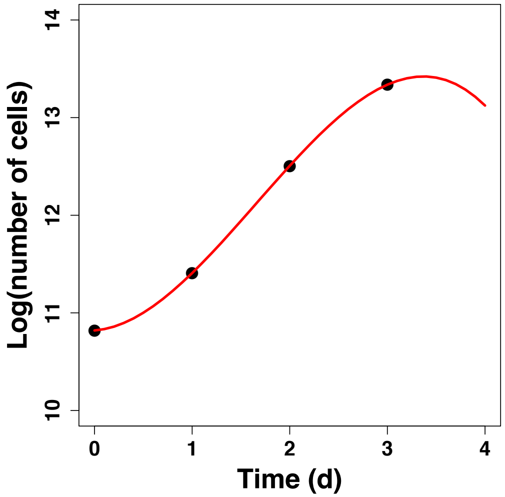
$$\begin{cases} y_1 = \theta_0 + \theta_1 t_1 + \theta_2 t_1^2 \\ y_2 = \theta_0 + \theta_1 t_2 + \theta_2 t_2^2 \\ y_3 = \theta_0 + \theta_1 t_3 + \theta_2 t_3^2 \end{cases} \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{pmatrix}$$

$$\Leftrightarrow y = M \cdot \theta \Leftrightarrow \theta = M^{-1} \cdot y$$

# Linear system: polynomial interpolation

- What if we have 3 points?
- 3 points ⇔ 3 degrees of freedom ⇔ 3 parameters

$$y = \theta_0 + \theta_1 t + \theta_2 t^2 + \theta_3 t^3$$




4 unknowns

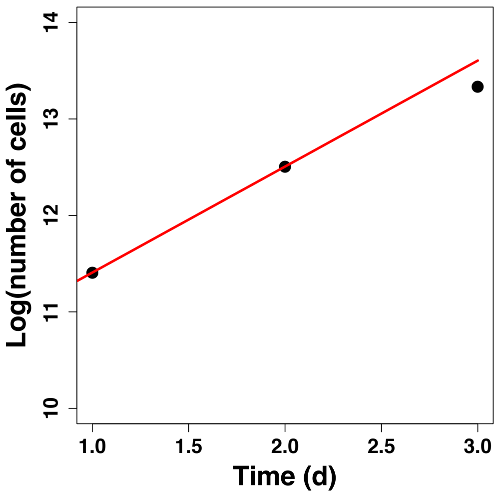
4 equations

$$\begin{cases} y_1 = \theta_0 + \theta_1 t_1 + \theta_2 t_1^2 + \theta_3 t_1^2 \\ y_2 = \theta_0 + \theta_1 t_2 + \theta_2 t_2^2 + \theta_3 t_2^2 \\ y_3 = \theta_0 + \theta_1 t_3 + \theta_2 t_3^2 + \theta_3 t_3^2 \\ y_4 = \theta_0 + \theta_1 t_3 + \theta_2 t_3^2 + \theta_3 t_4^2 \end{cases} \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & t_1 & t_1^2 \\ 1 & 1 & t_2 & t_2^2 \\ 1 & 1 & t_3 & t_3^2 \\ 1 & 1 & t_4 & t_4^2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

⇒ overfit, poor predictive power

# Back to simplicity: line

- How to fit 3 points with one line?




2 unknowns

3 equations

$$\begin{cases} y_1 = \theta_0 + \theta_1 t_1 \\ y_2 = \theta_0 + \theta_1 t_2 \\ y_3 = \theta_0 + \theta_1 t_3 \end{cases} \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \theta_0 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \theta_1 \cdot \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

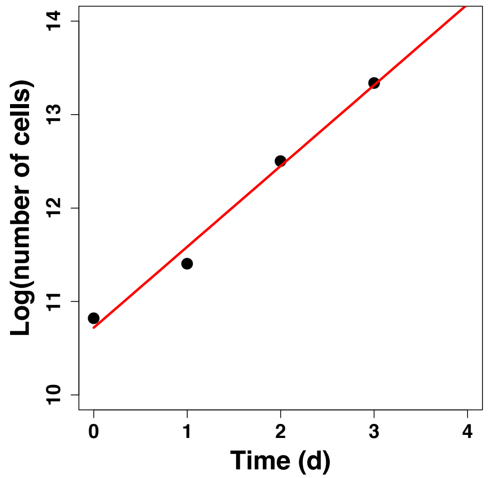
2 vectors cannot span a space of dimension 3

no solution (in general)

# Linear regression

$$y = \theta_0 + \theta_1 t + \varepsilon$$

Question: what is the « best » linear approximation of $y$ ?




2

$$n \quad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix} \longrightarrow$$

$M$ rectangular

no solution

$$\Leftrightarrow y = M \cdot \theta$$

$$\times\ M^T (\in M_{2,n}) \qquad \Rightarrow M^T y = M^T M \cdot \theta \longrightarrow$$

one unique solution

(if the square matrix $M^T M$ is invertible)

$$\underbrace{M^T y}_{M_{2,n} \cdot M_{n,1}} \qquad \underbrace{M^T M \cdot \theta}_{M_{2,n} \cdot M_{n,2} \cdot M_{2,1}}$$

$$M_{2,1} \qquad M_{2,2} \cdot M_{2,1}$$

$$\hat{\theta} = \left(M^T M\right)^{-1} M^T y$$

# Linear least-squares

- $\hat{\theta}$ is the value of the parameter vector $\theta$ that minimizes the sum of squared residuals

$$SS = \sum_{i=1}^{n} \left(y_i - (\theta_0 + \theta_1 t_i)\right)^2 \qquad \hat{\theta}_1 = \frac{\sum (y_i - \bar{y})(t_i - \bar{t})}{\sum (t_i - \bar{t})^2}, \qquad \hat{\theta}_0 = \bar{y} - \hat{\theta}_1 \bar{t}$$

- It is called the least-squares estimator of the linear model

- It corresponds to the projection of $y \in \mathbb{R}^n$ on the column space of the matrix $M$, i.e the space spanned by $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and $t = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$, of dimension 2 (2 linearly independent vectors)

- It regresses the information contained in the dependent variable $y$ on the independent variables $\mathbf{1}$ (constants) and $t$

# 1.2 General theory

# Formalism

- **Observations**: *n* couples of points $(t_j, y_j)$, with $y_j \in \mathbb{R}$ (or $\mathbb{R}^m$).

  We will denote $y = (y_1, \cdots, y_n) \in \mathbb{R}^n$ and $t = (t_1, \cdots, t_n)$.

- **Structural model:** a function

$$M : \begin{array}{ccc} \mathbb{R} \times \mathbb{R}^p & \to & \mathbb{R} \\ (t, \theta) & \mapsto & M(t, \theta) \end{array}$$

- The (unknown) vector of **parameters** $\theta \in \mathbb{R}^p$

**Goal = find $\theta$**

# Statistical model

$$y_j = M(t_j; \theta^*) + e_j$$

- « True » parameter $\theta^*$
- $e_j$ = **error** = measurement error + structural error
- **Random variables,** often independent and identically distributed

$$Y_j = M(t_j; \theta^*) + \varepsilon_j$$

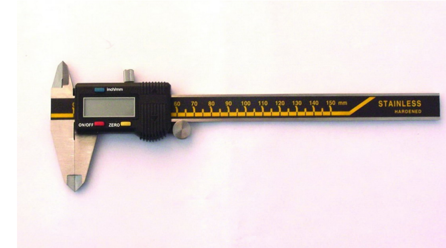
$Y_j, \varepsilon_j$ = r.v.

$y_j, e_j$ = realizations

- $(y_1, \cdots, y_n)$ = **sample** with probability density function $p(y \mid \theta^*)$
- An **estimator** of $\theta^*$ is a random variable function of $Y$, denoted $\hat{\theta}$:

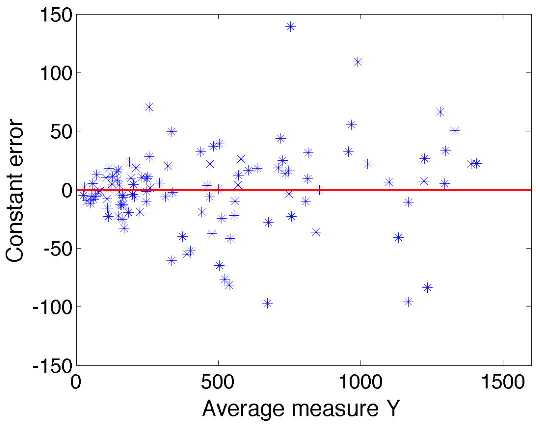
$$\hat{\theta} = h(Y_1, \cdots, Y_n)$$

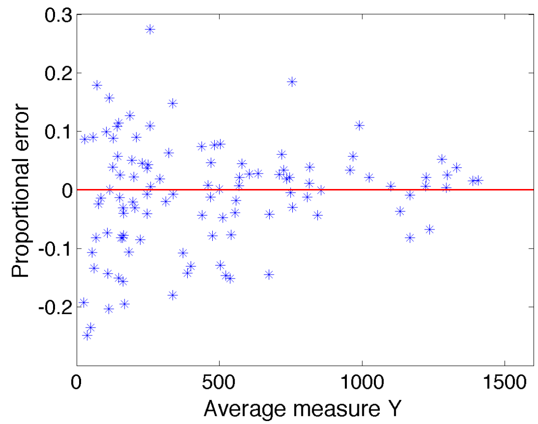
# Error models for tumor volume

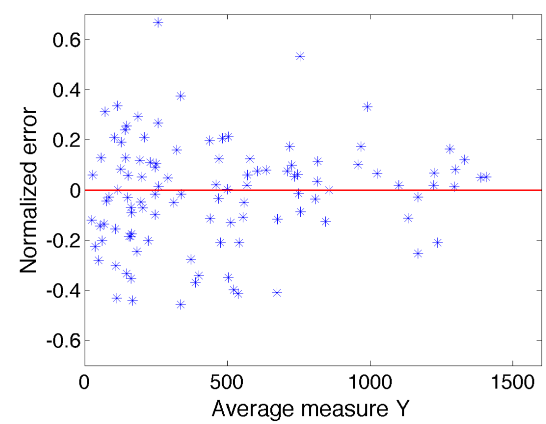


$$\varepsilon_j \text{ i.i.d } \mathcal{N}(0, \sigma_j)$$

| **Constant** | **Proportional** | **Specific** |
| --- | --- | --- |
| $\sigma_j = \sigma, \forall j$ | $\sigma_j = \sigma M\left(t_j, \hat{\theta}\right)$ | $\sigma_i = \begin{cases} \sigma M\left(t_j, \hat{\theta}\right)^{\alpha}, & M\left(t_j, \hat{\theta}\right) \geq V_m \\ \sigma V_m^{\alpha}, & M\left(t_j, \hat{\theta}\right) < V_m \end{cases}$ |
| $p = 0.004$ | $p = 0.083$ | $p = 0.2$ |



# Linear least-squares: statistical properties

$$Y = M\theta^* + \varepsilon$$

$$\hat{\theta}_{LS} = \underset{\theta \in \mathbb{R}^p}{\text{argmin}} \; \| Y - M\theta \|^2 \Leftrightarrow \hat{\theta}_{LS} = \left(M^T M\right)^{-1} M^T Y$$

***Proposition:***

*Assume that* $\varepsilon \sim \mathcal{N}\left(0,\sigma^2 I\right)$*, then*

$$\hat{\theta}_{LS} \sim \mathcal{N}\left(\hat{\theta}^*, \sigma^2 \left(M^T M\right)^{-1}\right)$$

From this, standard errors and confidence intervals can be computed on the parameter estimates

$$se\left(\hat{\theta}_{LS,p}\right) = \sigma \sqrt{\left(M^T M\right)^{-1}_{p,p}}$$

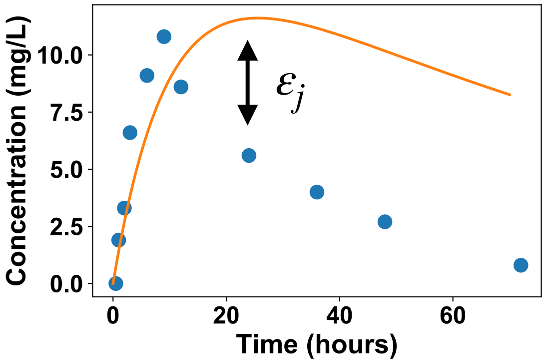
$$IC_\alpha\left(\theta_{LS,p}\right) = \theta^* \pm t^{\alpha/2}_{n-p} s \sqrt{\left(M^T M\right)^{-1}_{p,p}}$$

$$s^2 = \frac{1}{n-p} \; \| y - M\hat{\theta}_{LS} \|^2$$

# Nonlinear regression: least-squares

$$Y = M(t; \theta^*) + \varepsilon$$

$$\hat{\theta}_{LS} = \underset{\theta \in \mathbb{R}^p}{\text{argmin}} \left\| Y - M(t; \theta) \right\|^2$$




Linearization: $M(t, \theta) = M(t, \theta^*) + J \cdot (\theta - \theta^*) + o(\theta - \theta^*), \quad J = D_\theta M(t, \theta^*)$

***Proposition:***

*Assume $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I\right)$. Then, for large $n$, approximately*

$$\hat{\theta}_{LS} \sim \mathcal{N}\left(\hat{\theta}^*, \sigma^2 \left(J^T J\right)^{-1}\right)$$

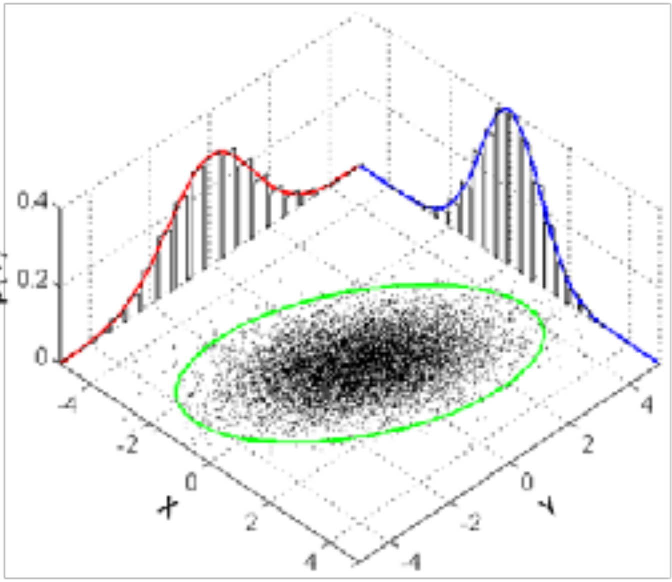
**$\Rightarrow$ standard errors, confidence intervals**

# Sensitivity matrix

$$J = D_\theta M(t,\hat{\theta}) = \begin{pmatrix} \frac{\partial M}{\partial \theta_1}\left(t_1,\hat{\theta}\right) & \cdots & \frac{\partial M}{\partial \theta_p}\left(t_1,\hat{\theta}\right) \\ \vdots & \ddots & \vdots \\ \frac{\partial M}{\partial \theta_1}\left(t_n,\hat{\theta}\right) & \cdots & \frac{\partial M}{\partial \theta_p}\left(t_n,\hat{\theta}\right) \end{pmatrix}$$

$$var\left(\hat{\theta}_{LS}\right) = \sigma^2\left(J^T J\right)^{-1}$$

- $J^T J$ is a $p \times p$ symmetric matrix
- It is invertible if and only if $rank(J) = p$
- Column $k$ of $J = 0 \Leftrightarrow M(t,\hat{\theta})$ does not depend on $\theta_k$
- Line $i$ of $J = 0 \Leftrightarrow M(t_i,\hat{\theta})$ does not depend on $\theta$



# Nonlinear regression: Likelihood maximization

$$Y = M(t; \theta^*) + \varepsilon$$

The likelihood is defined by

$$L(\theta) = p(y_1, \cdots, y_n | \theta) = \prod_{j=1}^{n} p(y_j | \theta)$$

It is the probability to observe $y$ if the parameter is $\theta$.

The maximum likelihood estimator (MLE) is the value of $\theta$ that maximizes the likelihood

$$\hat{\theta}_{MV} = \underset{\theta}{\text{argmax}} L(\theta)$$

# Asymptotic properties of the MLE

***Proposition:***

*Under regularity assumptions on $L$, when $n \to +\infty$*

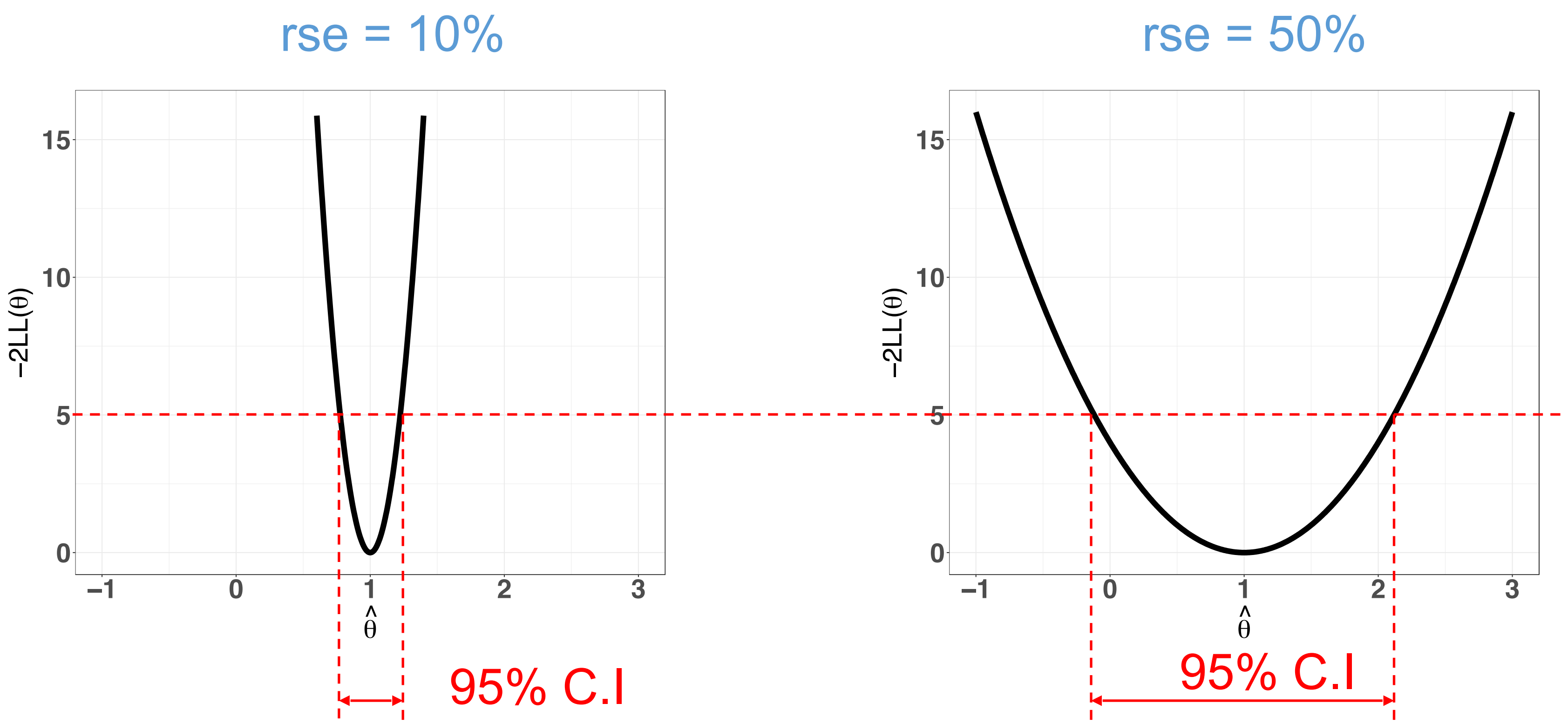
1. $\hat{\theta}_{MV} \longrightarrow \theta^*$ *(consistency)*
2. $\hat{\theta}_{MV}$ *is asymptotically of minimal variance (it reaches the Cramér-Rao bound):*

$$\sqrt{n}\left(\hat{\theta}_{MV} - \theta^*\right) \rightharpoonup \mathcal{N}\left(0, I_{\theta^*}^{-1}\right)$$

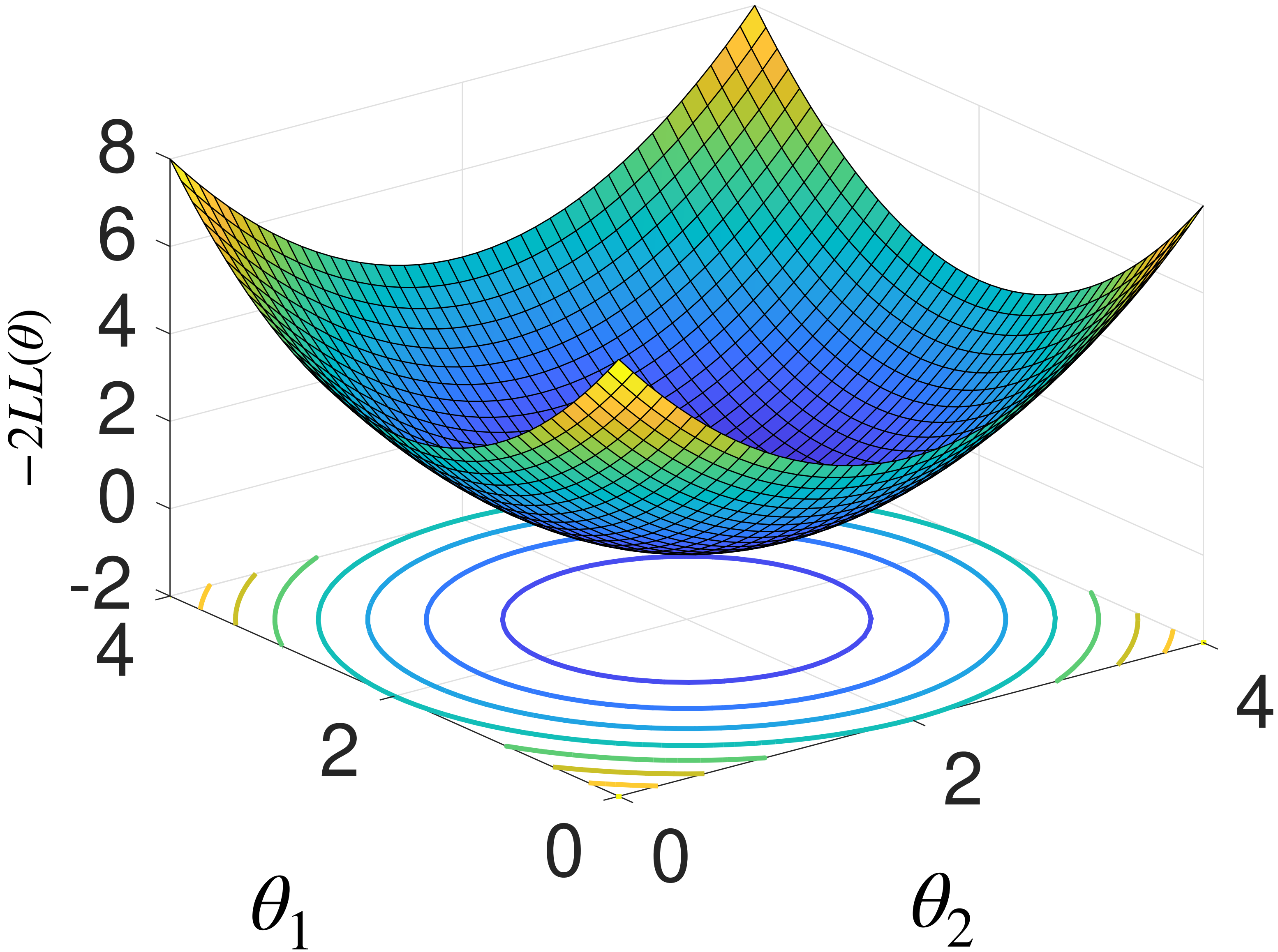
*where $I_{\theta^*}$ is the Fisher information matrix*

$$\left(I_{\theta^*}\right)_{j,k} = \mathbb{E}\left[\left\{\frac{\partial \log(p(Y|\theta^*))}{\partial \theta_j}\right\}\left\{\frac{\partial \log(p(Y|\theta^*))}{\partial \theta_k}\right\}\right] = \mathbb{E}\left[-\left(\frac{\partial^2 \log\left(p\left(Y|\theta^*\right)\right)}{\partial \theta_j \partial \theta_k}\right)\right].$$

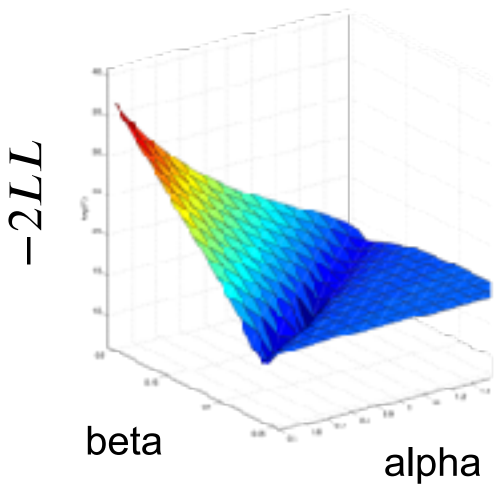
# Precision of the estimates

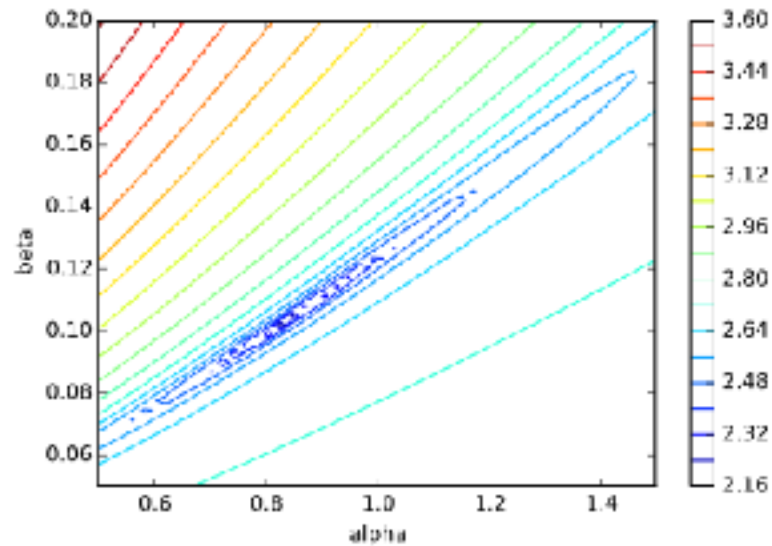


## In 2D



# Correlation between estimates




Correlation matrix of the estimates

| | R.S.E.(%) | | | |
|---|---|---|---|---|
| alpha_pop | 3.09 | 1 | | |
| beta_pop | 5.65 | 0.96874 | 1 | |
| b | 23.3 | -0.0055974 | 0.002018 | 1 |

| | MIN | MAX | MAX/MIN |
|---|---|---|---|
| Eigen values | 0.014 | 2 | 1.4e+2 |

small r.s.e on alpha and beta, but large correlation

# MLE: normal errors

$$Y_j = M\left(t_j; \theta^*\right) + \varepsilon_j, \quad \varepsilon_j \sim \mathcal{N}(0, \sigma)$$

$$p(y_j \mid \theta, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left(y_j - M(t_j, \theta)\right)^2}{2\sigma^2}}, \quad L(\theta, \sigma) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} e^{-\frac{\| y - M(t, \theta) \|^2}{2\sigma^2}}$$

Maximize $L(\theta, \sigma)$ $\Leftrightarrow$ minimize $F(\theta, \sigma) = -\log\left(L(\theta, \sigma)\right)$

$$F(\theta, \sigma) = n \log\left(\sigma\sqrt{2\pi}\right) + \frac{\| y - M(t, \theta) \|^2}{2\sigma^2}$$

$$\frac{\partial F}{\partial \sigma}\left(\hat{\theta}, \hat{\sigma}\right) = 0 \Rightarrow \hat{\sigma} = \frac{1}{n} \| y - M(t, \hat{\theta}) \|^2$$

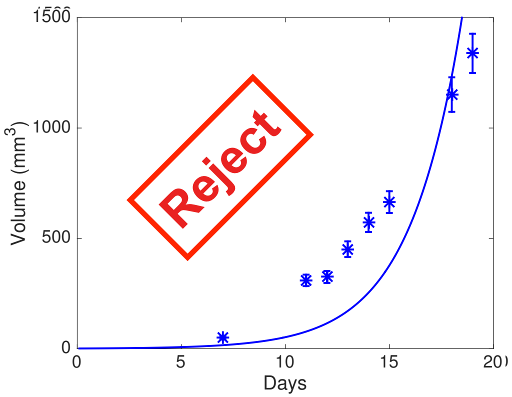
$$\Rightarrow \hat{\theta} = \underset{\theta}{\text{argmin}} \| y - M(t, \theta) \|^2$$

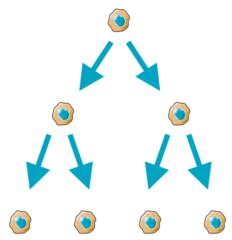
**Maximum likelihood ⇔ Least-squares**

# Application: tumor growth

What are **minimal** biological processes able to recover the **kinetics** of (experimental) tumor growth?

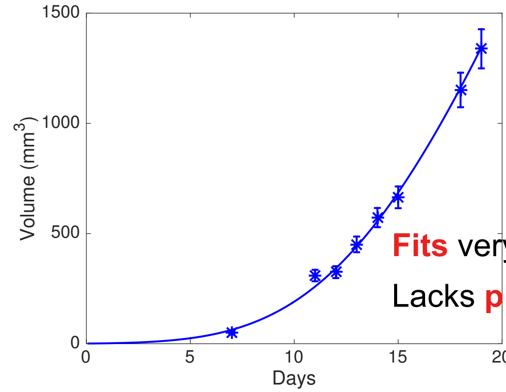
## Exponential



$$\frac{dV}{dt} = aV$$

## Gompertz

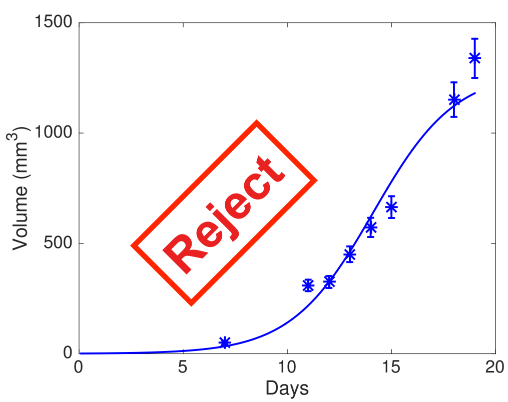



$$\frac{dV}{dt} = \alpha e^{-\beta t} V$$

**Fits** very well

Lacks **physiological** interpretation

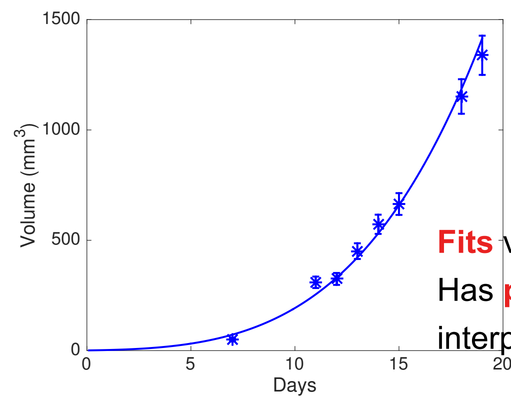
## Logistic




Competition

$$\frac{dV}{dt} = aV\left(1 - \frac{V}{K}\right)$$

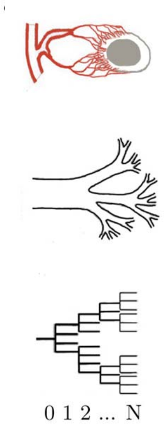
## Power law




$$\frac{dV}{dt} = \alpha V^{\gamma}$$

**Fits** very well

Has **physiological** interpretation




*Benzekry et al., PloS Comp Biol, 2014*

# Goodness of fit metrics

## Sum of Squared Errors

$$SSE^i = \sum_{j=1}^{n^i} \left( \frac{V_j^i - V(t_j^i, \hat{\theta}^i)}{\hat{\sigma}_j^i} \right)^2$$

## Akaike Information Criterion

$$AIC^i = -2l(\hat{\theta}^i) + 2p$$

↑ number of parameters

| **Model** | **SSE** | **AIC** | **RMSE** | **R2** | **p > 0.05** | **#** |
|---|---|---|---|---|---|---|
| Power law | **0.164(0.0158 - 0.646)[1]** | **-18.4(-43.2 - 1.63)[1]** | **0.415(0.145 - 0.899)[1]** | **0.97(0.801 - 0.998)[1]** | 100 | 2 |
| Gompertz | 0.176(0.019 - 0.613)[2] | -16.9(-48.2 - 1.1)[2] | 0.433(0.156 - 0.875)[2] | 0.971(0.828 - 0.997)[2] | 100 | 2 |
| Logistic | 0.404(0.0869 - 0.85)[3] | -5.41(-18.4 - 3.88)[3] | 0.665(0.331 - 1)[3] | 0.908(0.712 - 0.989)[3] | 100 | 2 |
| Exponential | 1.9(0.31 - 3.56)[4] | 10.7(-5.38 - 23.1)[4] | 1.4(0.595 - 1.95)[4] | 0.69(0.454 - 0.944)[4] | 15 | 1 |

## Root Mean Squared Errors

$$RMSE^i = \sqrt{\frac{1}{n-p} SSE^i}$$

## $R^2$

$$R^{2,j} = 1 - \frac{\sum_j \left( V_j^i - V(t_j^i; \hat{\theta}^i) \right)^2}{\sum_j \left( V_j^i - \overline{V^i} \right)^2}$$

# Parameter values and identifiability

| Model | Par. | Unit | Median value (CV) | NSE (%) (CV) |
| --- | --- | --- | --- | --- |
| Power law | $\alpha$ | $mm^{3(1-\gamma)} \cdot day^{-1}$ | 0.886 (30.8) | 8.17 (52.5) |
| | $\gamma$ | - | 0.788 (7.56) | 2.28 (58.6) |
| Gompertz | $\alpha_0$ | $day^{-1}$ | 1.68 (23.5) | 6.11 (82.9) |
| | $\beta$ | $day^{-1}$ | 0.0703 (28) | 8.35 (92.9) |
| Logistic | $a$ | $day^{-1}$ | 0.474 (13.3) | 2.93 (23.3) |
| | $K$ | $mm^3$ | 1.92e+03 (36.7) | 15.8 (28.7) |
| Exponential | $a$ | $day^{-1}$ | 0.356 (12.9) | 2.53 (19.4) |
| Generalized logistic | $a$ | $[day^{-1}]$ | 2555 (148) | 2.36e+05 (137) |
| | $K$ | $[mm^3]$ | 4378 (307) | 165 (220) |
| | $\alpha$ | - | 0.0001413 (199) | 2.36e+05 (137) |

NSE = Normalized Standard Error ⟷ practical identifiability

$$\hat{\theta} \sim \mathcal{N}\left(\theta^*, \hat{\sigma}^2 \left(J \cdot J^T\right)^{-1}\right)$$

$$se\left(\hat{\theta}^k\right) = \sqrt{\hat{\sigma}^2 \left(J \cdot J^T\right)_{k,k}}$$

# References


- Course « Statistics in Action with R » by Marc Lavielle
  http://sia.webpopix.org/index.html
- Seber, G. A., & Wild, C. J. (2003). Nonlinear regression. Hoboken (NJ): Wiley-Interscience.