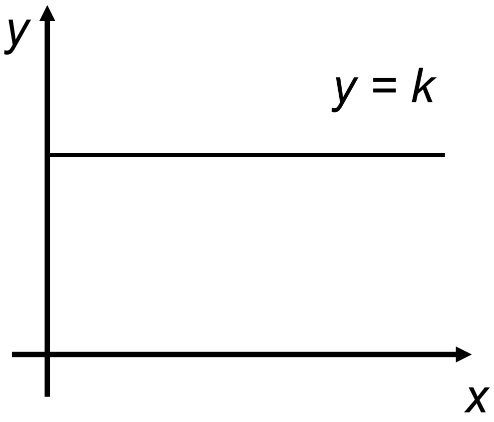
Mathematical tools for pharmacometrics: Linear algebra

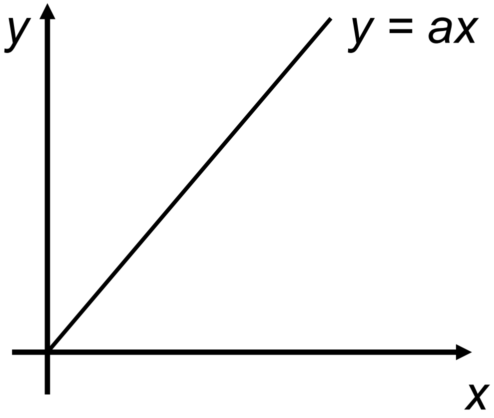
S. Benzekry

sebastien.benzekry@inria.fr

Linearity: the simplest mathematical relationship

- Variable y (ex: concentration max) that depends on variable x (ex: dose)





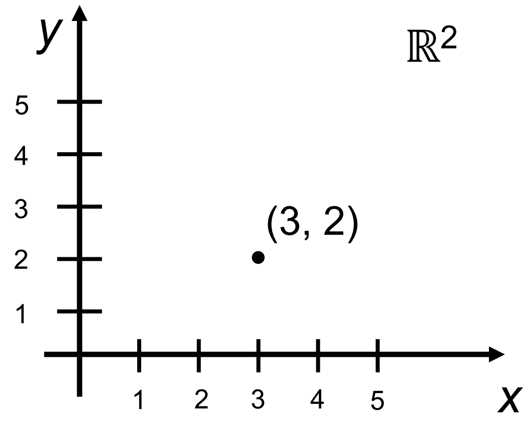
How to extend linearity to several variables $(x_1, \ldots, x_n)$? (ex: dose, weight,…)

Multiple variables = vectors

- a single (real) number is called a scalar : 1, -2, 3.5, 5/7, π, etc…
- an ordered set of numbers is called a vector: (x_1, x_2, x_3), (2, -1, 5) ≠ (-1, 2, 5)
- mathematically, it's an element of $\mathbb{R}^n$, n is called the dimension
- a vector can be written as row or column

(2, -1, 5)

$$\begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}$$



Examples

- All covariates of one patient

SEX	AGE	WEIGHT	BSA	CREA	CREA_CLR	HB	PQ	PNN	ASAT	ALAT	ALB
1	64	76	2	53	145	106	225	10.4	93	101	25.6

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}

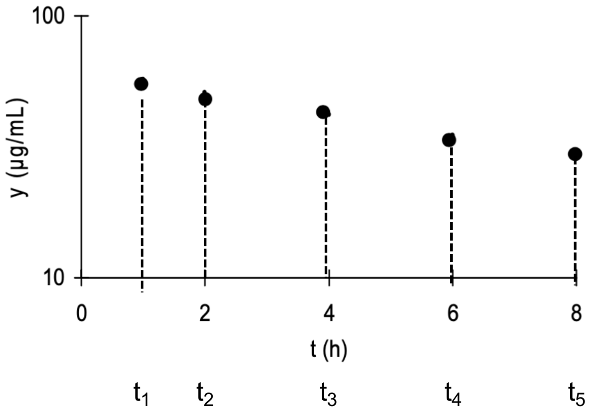
- One covariate in all patients

WEIGHT	
76	x^1
57	x^2
62	
77	
82	
77	
81	
83	
93	
61	
78	
74	
74	
71	
45	
79	
64	
63	
67	
72	
88	x^n

- Sampling times and observations

$(t_1, \ldots, t_n)$

$(y_1, \ldots, y_n)$



Operations on vectors

- Addition of two vectors: $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$

  $$(2, -1, 5) + (-1, 3, 2) = (1, 2, 7)$$

- Multiplication of two vectors? $(x_1, \ldots, x_n) * (y_1, \ldots, y_n) = (x_1\, y_1, \ldots, x_n\, y_n)$?

- Multiplication by a scalar: $\lambda\,(x_1, \ldots, x_n) = (\lambda\, x_1, \ldots, \lambda\, x_n)$. $3\,(2, -1, 5) = (6, -3, 15)$

$$\Rightarrow \text{vector space } \mathbb{R}^n$$

$$f: \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R} \\ x = (x_1, \cdots, x_n) & \mapsto & f(x) = f\,(x_1, \cdots, x_n) \end{array}$$

f linear?

$$f(x+y) = f(x) + f(y) \quad f(\lambda x) = \lambda f(x)$$

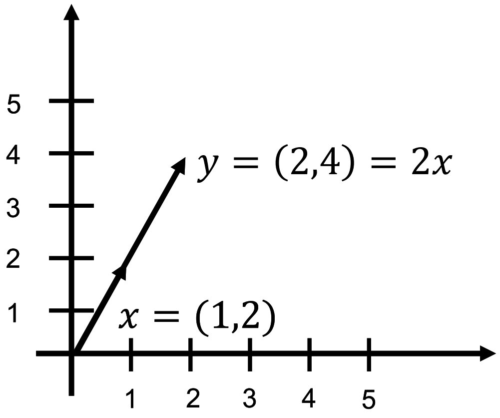
$$\Rightarrow \exists\, a = (a_1, \cdots, a_n)\,,\ f(x) = a_1 x_1 + \cdots + a_n x_n := \langle a, x \rangle$$

Linear combinations

- If two vectors x and y are such that $y = \lambda x$ for $\lambda \in \mathbb{R}$, y and x are colinear

  $\Rightarrow$ they carry the same information

x	y
Fluorescence	Bioluminescence
8.53E+09	1.50E+08
2.19E+10	3.33E+08
5.20E+09	2.25E+08
1.41E+10	2.62E+08
8.05E+09	4.72E+08
9.51E+09	1.79E+09
1.06E+10	9.03E+08

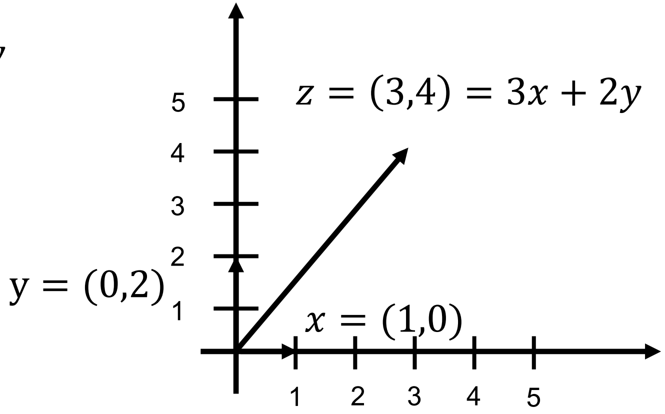


- Linear combination of two vectors x and y: $z = \lambda x + \mu y$

BSA = 0.00718 x $W^{0.425}$ x $H^{0.725}$

ln(BSA) = ln(0.00718) + 0.425 ln(W) + 0.725 ln(H)

$\Rightarrow$ the three quantities W, H and BSA are redundant



Basis

- *n* vectors $\{x_1, \cdots, x_n\}$ are said to be linearly independent if no vector can be expressed as a linear combination of the others

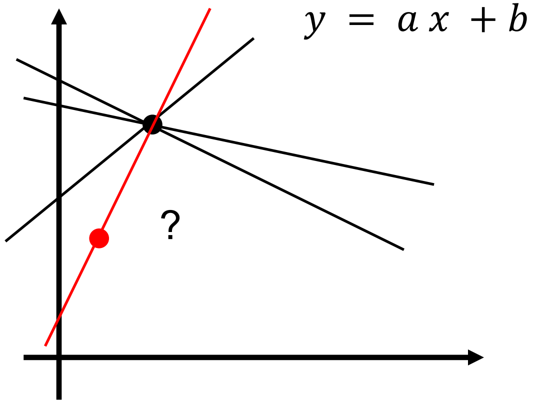
- Thm: In dimension *n* any vectors $\{x_1, \cdots, x_p\}$ with *p* > *n* are linearly dependent

$\Rightarrow$ with *n* observations it is only possible to identify maximum *n* parameters

- A set $\{x_1, \cdots, x_p\}$ of vectors is said to span $\mathbb{R}^n$ if any vector of $\mathbb{R}^n$ is a linear combination of $\{x_1, \cdots, x_p\}$:

$$\forall x \in \mathbb{R}^n, \exists\, \lambda_1, \cdots, \lambda_p \in \mathbb{R}\ \ s.t.\ x = \lambda_1 x_1 + \cdots + \lambda_p x_p$$

- Thm: In dimension *n* any set of *n* linearly independent vectors spans $\mathbb{R}^n$ and is called a basis



Matrices

- A matrix is a rectangular array of numbers with a given number of rows (m) and columns (n)

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \vdots \\ m \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & \dots & n \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{array} = (a_{i,j})$$

3 columns

$$\begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix}$$

2 rows

Elementary operations

- Similarly as vectors, we can define addition and multiplication by a scalar

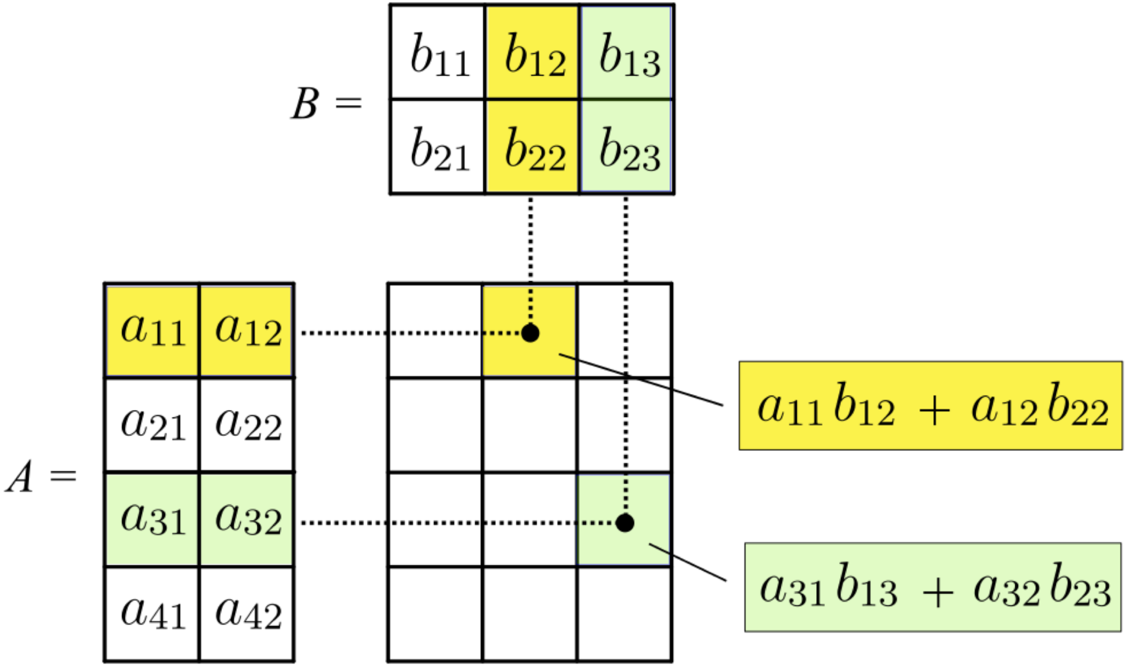
$$\begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 1 \\ 3 & -4 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 6 \\ 1 & 8 & -2 \end{pmatrix} \qquad 2 \cdot \begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix} = \begin{pmatrix} 6 & -2 & 10 \\ -4 & 24 & -14 \end{pmatrix}$$

$\Rightarrow$ the set of m-by-n matrices is a vector space $M_{m,n}$ of dimension $m.n$

- Transposition

$$\begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix}^T = \begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix}' = \begin{pmatrix} 3 & 2 \\ -1 & 12 \\ 5 & -7 \end{pmatrix}$$

Matrix multiplication



Matrix multiplication

$$\begin{pmatrix} 3 & -1 & 5 \\ -2 & 12 & -7 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 \times 3 + 0 \times (-1) + 2 \times 5 & \cdots \\ \cdots & \cdots \end{pmatrix} = \begin{pmatrix} 23 & \cdots \\ \cdots & \cdots \end{pmatrix}$$

2 rows, 3 columns (first matrix)

3 rows, 2 columns (second matrix)

2 rows, 2 columns (result)

- The number of rows of the second matrix must be the same as the number of columns of the first matrix
- It is only possible to multiply a matrix in $M_{m,n}$ by a matrix in $M_{n,p}$ and it gives a $M_{m,p}$ matrix

Special matrices

- Square matrix $\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$

- Symmetric matrix: $M^T = M$ $\begin{pmatrix} 1 & 5 & -3 \\ 5 & 2 & -1 \\ -3 & -1 & -4 \end{pmatrix}$

- Diagonal matrix $\begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$

- Identity matrix $I = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$

$$\forall M \in M_{n,n}, \qquad M \cdot I = I \cdot M = M$$

Example 1: data array

n columns (covariates)

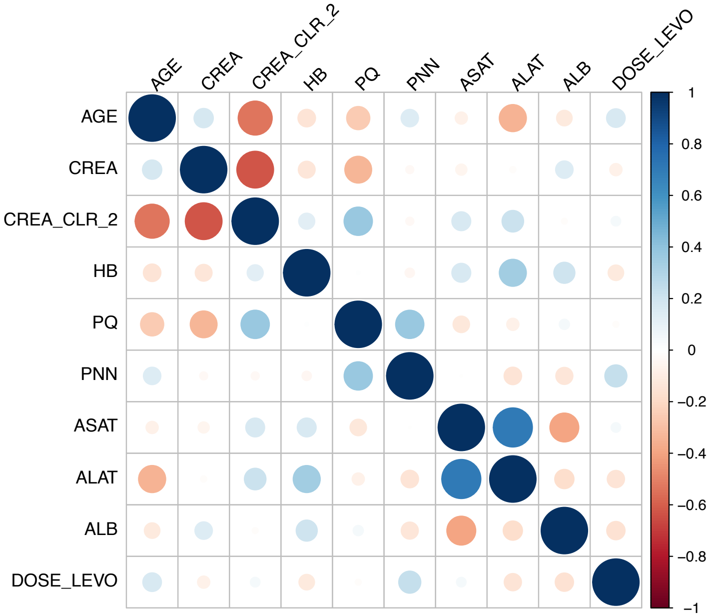
m patients

AGE	BSA	CREA	CREA_CLR_2	HB	PQ	PNN	ASAT	ALAT	ALB	DOSE_LEVO
64	2	53	140	106	225	10.4	93	101	25.6	450
37	1.5	37	195	152	185	5.75	25	70	44	300
61	1.75	30	200	112	428	24.71	34	32	34.8	5600
49	2	71.6	120	86	138	14.06	15	23	38.1	1250
49	2	56	136	78	103	4.03	30	67	36.1	0
33	1.79	54	211	79	441	16.95	22	34	36.6	0
33	1.79	71	152	119	300	5.42	30	54	34.3	0
33	1.79	52	177	105	146	4.86	21	34	34.3	0
36	1.87	67	128	129	279	4.76	45.5	118	39.1	0
47	2	56	149	85	400	8.76	15	15	35.4	3450
47	2	77	130	101	501	8.76	12	19	35.2	0
47	2	73	154	99	202	6.92	82	16	35.4	3140
59	1.83	71	125	86	22	0.2	35	85	37.4	2700

Example 2: correlation matrix

$(x_1, \cdots, x_n)$ n vectors, $C_{i,j} = corr(x_i, x_j)$

	AGE	CREA	CREA_CLR_2	HB	PQ	PNN	ASAT	ALAT	ALB	DOSE_LEVO
AGE	1.00	0.18	-0.54	-0.15	-0.25	0.14	-0.07	-0.35	-0.11	0.16
CREA	0.18	1.00	-0.63	-0.13	-0.34	-0.03	-0.06	-0.02	0.14	-0.07
CREA_CLR_2	-0.54	-0.63	1.00	0.12	0.38	-0.03	0.17	0.22	-0.02	0.04
HB	-0.15	-0.13	0.12	1.00	0.01	-0.04	0.17	0.34	0.21	-0.11
PQ	-0.25	-0.34	0.38	0.01	1.00	0.38	-0.13	-0.07	0.05	-0.02
PNN	0.14	-0.03	-0.03	-0.04	0.38	1.00	0.00	-0.15	-0.14	0.23
ASAT	-0.07	-0.06	0.17	0.17	-0.13	0.00	1.00	0.71	-0.39	0.04
ALAT	-0.35	-0.02	0.22	0.34	-0.07	-0.15	0.71	1.00	-0.18	-0.14
ALB	-0.11	0.14	-0.02	0.21	0.05	-0.14	-0.39	-0.18	1.00	-0.16
DOSE_LEVO	0.16	-0.07	0.04	-0.11	-0.02	0.23	0.04	-0.14	-0.16	1.00

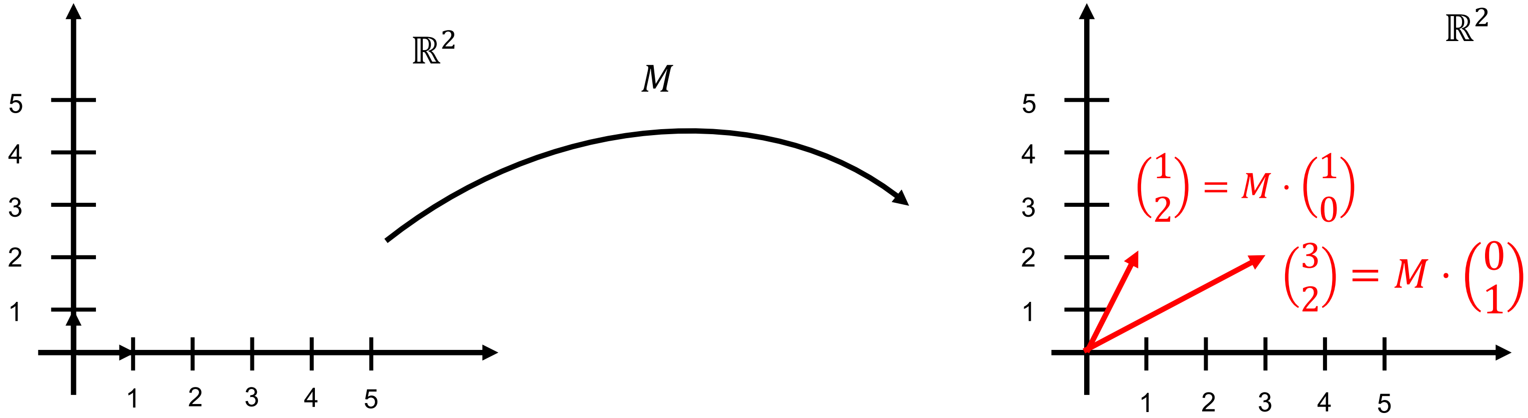


Linear map

- The product of a matrix $\mathrm{M} \in M_{n,n}$ with a vector $x \in \mathbb{R}^n$ gives a vector in $\mathbb{R}^n$.

$$\begin{pmatrix} 1 & 3 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

- $M: \mathbb{R}^n \to \mathbb{R}^n$ is linear: $M \cdot (x + y) = M \cdot x + M \cdot y,\ M \cdot (\lambda x) = \lambda M \cdot x$.



Linear system: Equation of a line

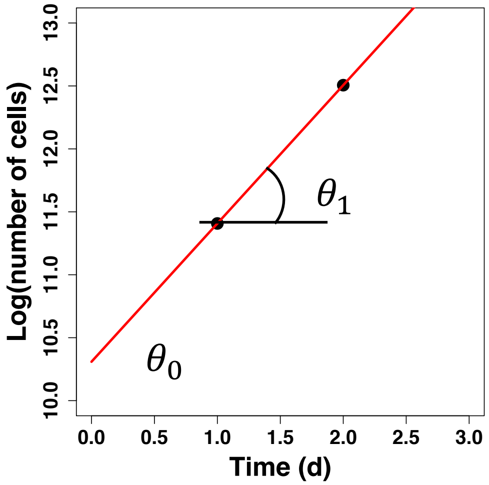
$$y = \theta_0 + \theta_1 t$$

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

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

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

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

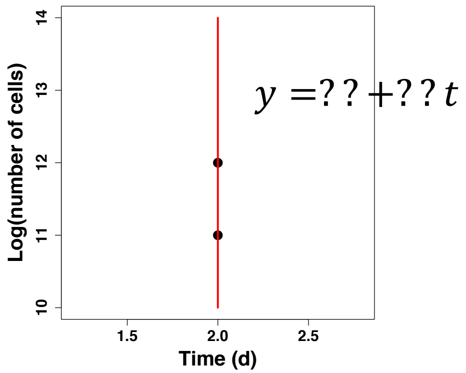


$$\begin{cases} 11.4 = 1\times\theta_0 + 1\times\theta_1 \\ 12.5 = 1\times\theta_0 + 2\times\theta_1 \end{cases} \Leftrightarrow \begin{pmatrix} 11.4 \\ 12.5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

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

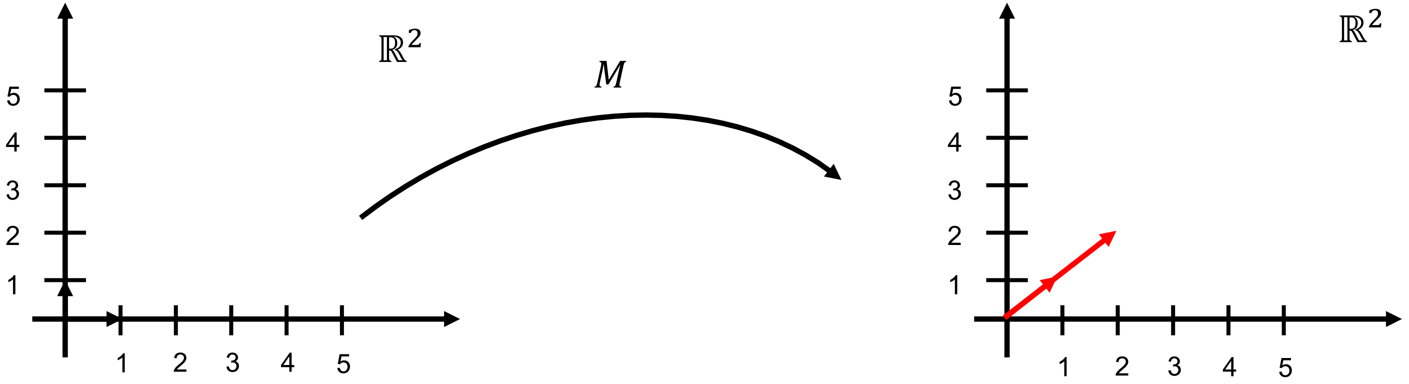
Doubling time = $\frac{\ln 2}{\theta_1} \times 24 = 15.1$ hours

Invertible matrix

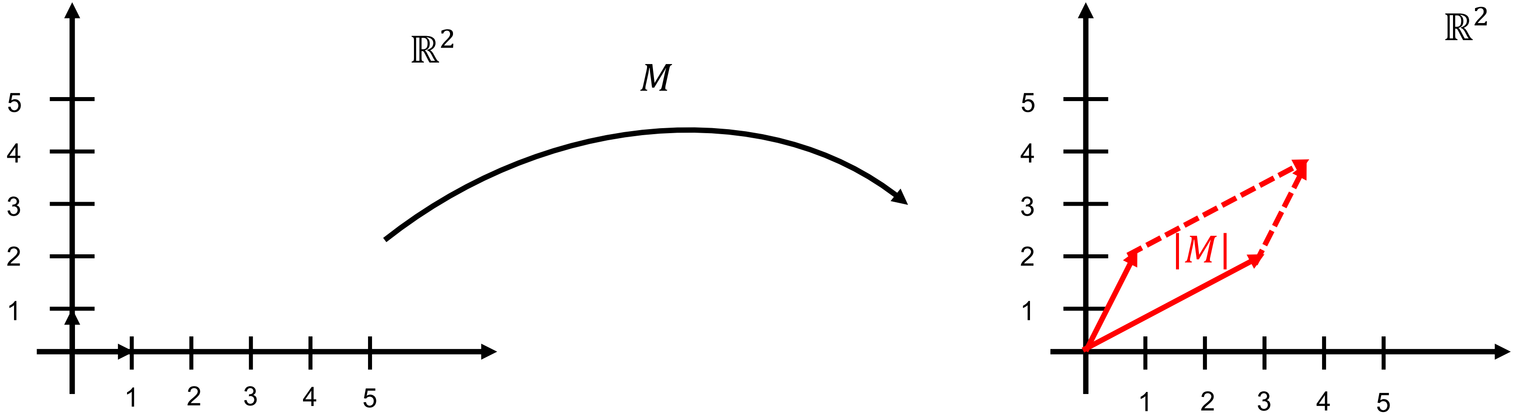


$$\begin{cases} 11 = \theta_0 + 2 \times \theta_1 \\ 12 = \theta_0 + 2 \times \theta_1 \end{cases} \Leftrightarrow \begin{pmatrix} 11 \\ 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

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



Determinant

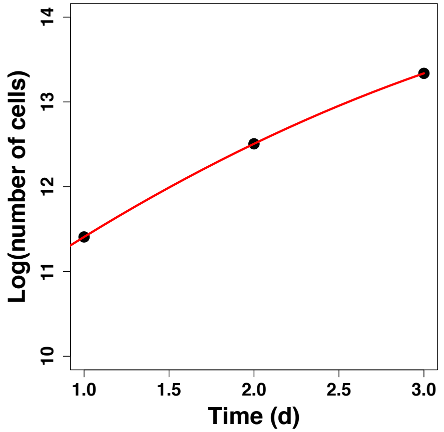


- The determinant of M, denoted $|M|$, is the area of the parallelogramm spanned by the column vectors of M
- For $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ it is given by $ad - bc$.
- It can be generalized in any dimension and is a measure of the colinearity (and correlation) of the vectors
- $|M| \neq 0 \Leftrightarrow$ 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

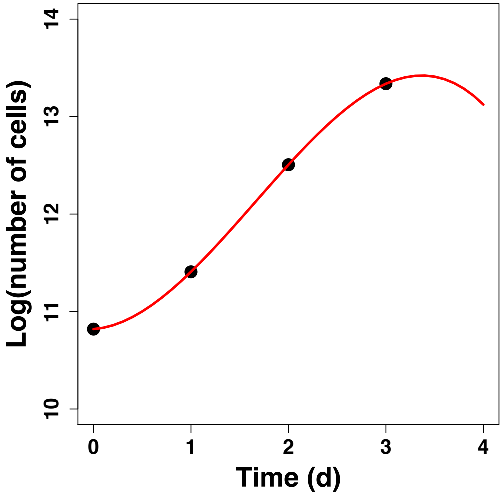
- 4 points?
- 4 points ⇔ 4 degrees of freedom ⇔ 4 parameters

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

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

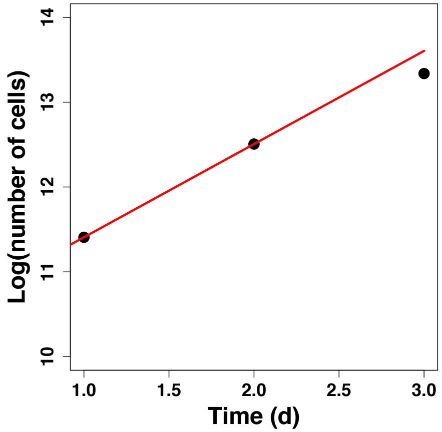
- How to fit 3 points with one line?

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

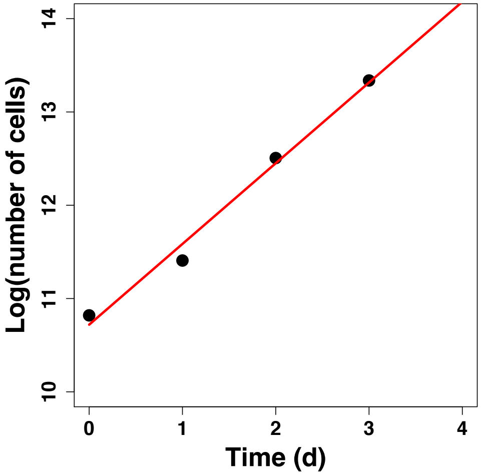
$$y = \theta_0 + \theta_1 t + \varepsilon$$

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

$$n \quad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{pmatrix}}^{2} \cdot \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}$$

⟶ M rectangular

no solution



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

$\times M^T (\in M_{2,n})$

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

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

⟶ one unique solution

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

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

Least-squares

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

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

- 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)

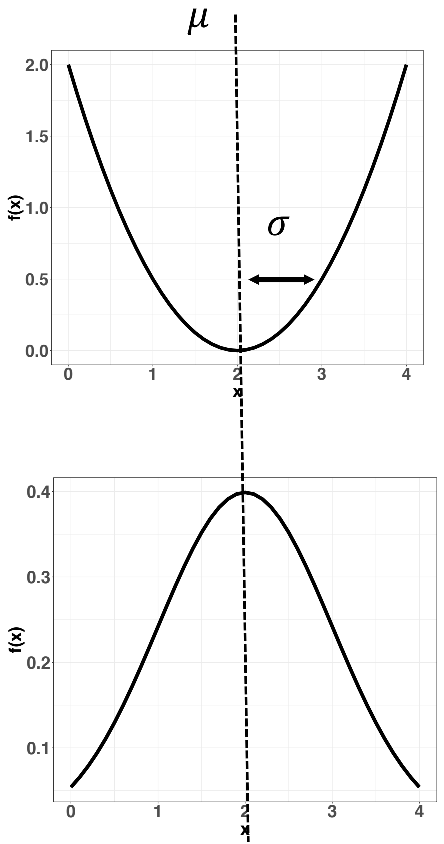
- It regresses the information contained in the dependent variable y on the independent variables $\mathbf{1}$ (constants) and t

Quadratic form: 1D. Normal distribution

- One complexity step beyond linearity: quadratic relationship

$$f: \begin{matrix} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & ax^2 \end{matrix}$$

$$f(x) = \frac{(x-\mu)^2}{2\sigma^2}$$



- Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

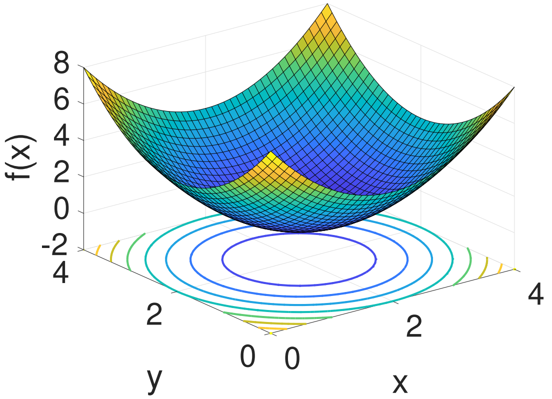
in 2D??

Quadratic form 2D: matrix form

- Quadratic form in $\mathbb{R}^2$

$$f : \begin{matrix} \mathbb{R}^2 & \rightarrow & \mathbb{R} \\ x = (x_1, x_2) & \mapsto & a x_1^2 + 2b x_1 x_2 + c x_2^2 \end{matrix}$$

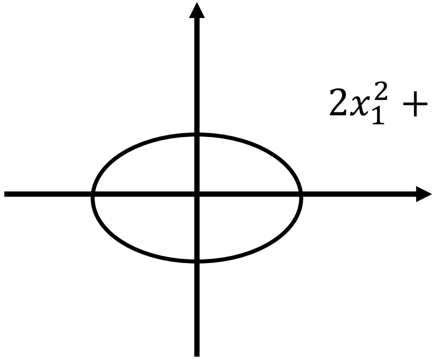
$$f(x) = (x_1, x_2) \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x^T \cdot M \cdot x$$



- M is a symmetric matrix

- If M is diagonal

$$(x_1, x_2) \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2x_1^2 + 3x_2^2$$



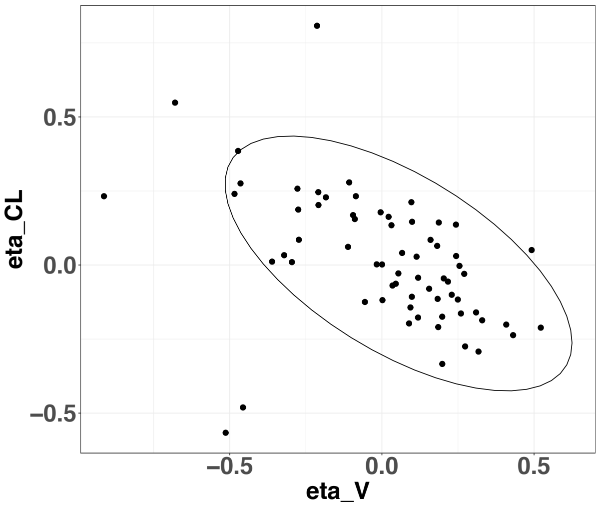
$2x_1^2 + 3x_2^2 = C$

Covariance/correlation matrix

- Two (or more) variables x and y (ex: V and CL)

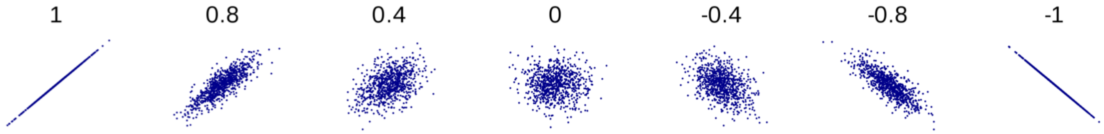
$$\Sigma = \begin{pmatrix} Var(x) & Cov(x,y) \\ Cov(x,y) & Var(y) \end{pmatrix}$$

$$Cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$



- Correlation matrix

$$R = \begin{pmatrix} 1 & r(x,y) \\ r(x,y) & 1 \end{pmatrix}$$



$$r(x,y) = \frac{Cov(x,y)}{\sigma_x \sigma_y}$$

note: $\widehat{\theta_1} = r(x,y)\frac{\sigma_y}{\sigma_x}$

Multivariate normal distribution

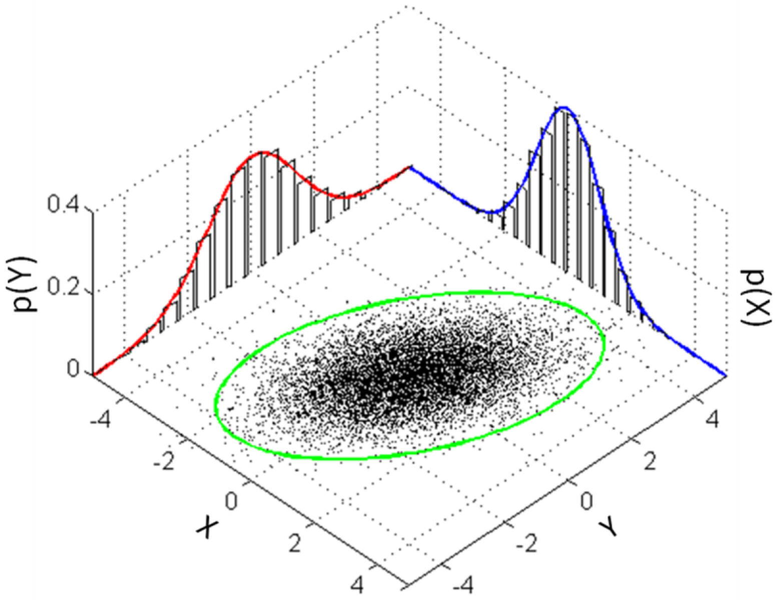
- Generalization of the normal distribution in dimension n

$$x \rightarrow \boldsymbol{x} = (x, y), \mu \rightarrow \boldsymbol{\mu} = (\mu_1, \mu_2)$$

$$\sigma^2 \rightarrow \Sigma = \begin{pmatrix} Var(x) & Cov(x,y) \\ Cov(x,y) & Var(y) \end{pmatrix}$$

$$\frac{(x-\mu)^2}{2\sigma^2} \rightarrow \frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})$$

$$f(x) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}$$



Eigenvalues and eigenvectors

- An eigenvector $v \in \mathbb{R}^n$ associated to an eigenvalue $\lambda \in \mathbb{R}$ is defined by

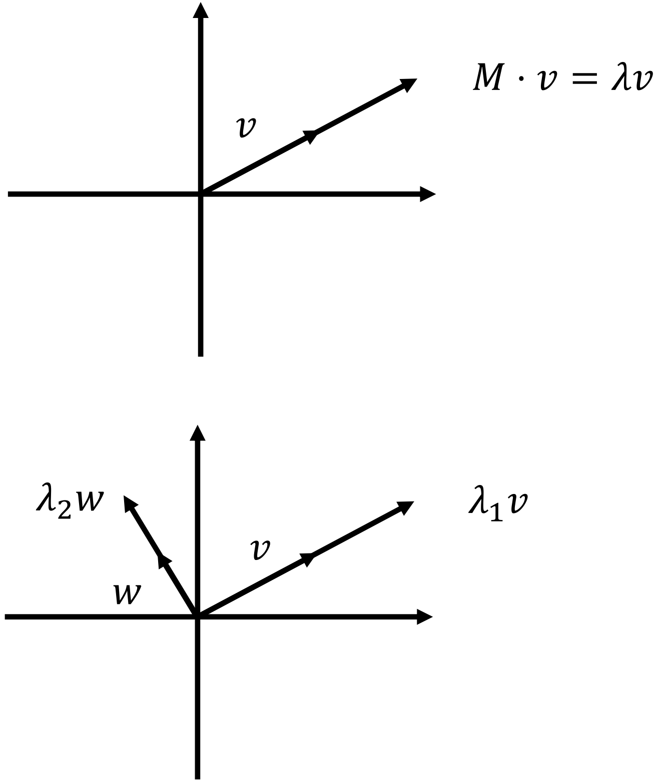
$$M \cdot v = \lambda v$$

- In a basis of eigenvectors, M is diagonal

$$M \triangleq D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

- Thm: If M is a symmetric matrix, it is diagonalizable (in an orthogonal basis)

$$M = PDP^{-1}, \qquad P^T P = I, \qquad P = (v, w)$$



Example

$$\Sigma = \begin{pmatrix} 2.98 & -4.65 \\ -4.65 & 64.1 \end{pmatrix}$$

$$\Sigma = P \cdot \begin{pmatrix} 64.4 & 0 \\ 0 & 2.63 \end{pmatrix} P^{-1}$$

- first eigenvector = direction of the data of maximal variance

- first eigenvalue = variance of the data in this direction