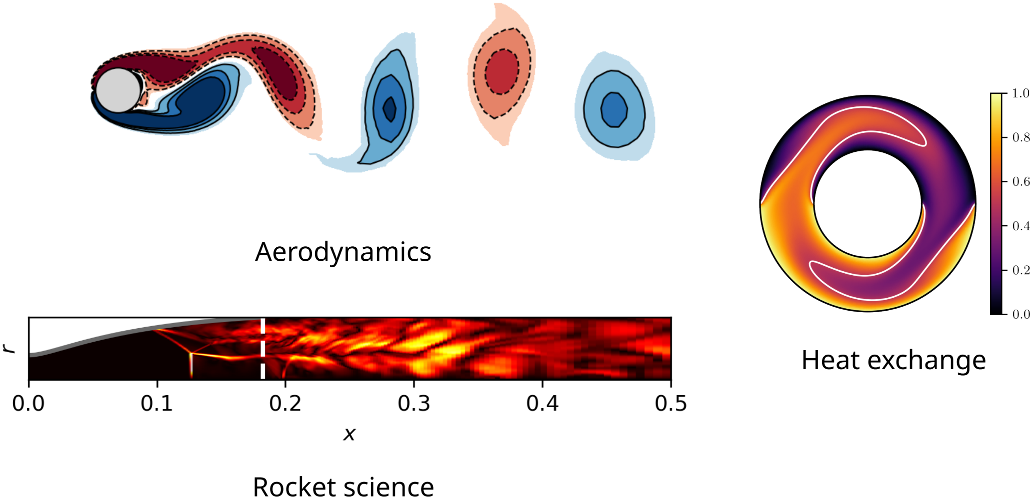
On the importance of low-dimensional structures for data-driven modeling

Jean-Christophe Loiseau

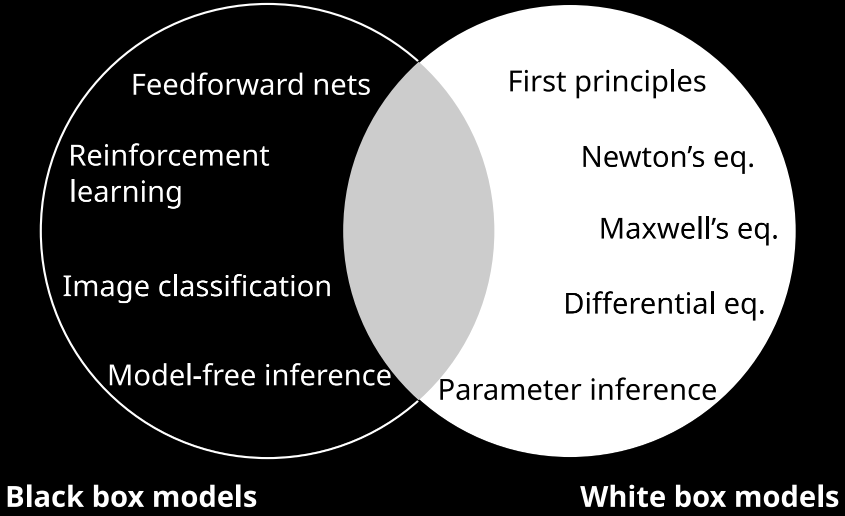
May, 13^{th} 2022

Who am I ?

- Maître de Conférences in Fluid Dynamics and Applied Math.
- Machine-learning enthusiast with application to engineering systems.
- Data-efficient models with guarantees of optimality or interpretability.


Aerodynamics
r
0.0
0.1
0.2
0.3
0.4
0.5
x
Rocket science
1.0
0.8
0.6
0.4
0.2
0.0
Heat exchange

N. Wiener's classification



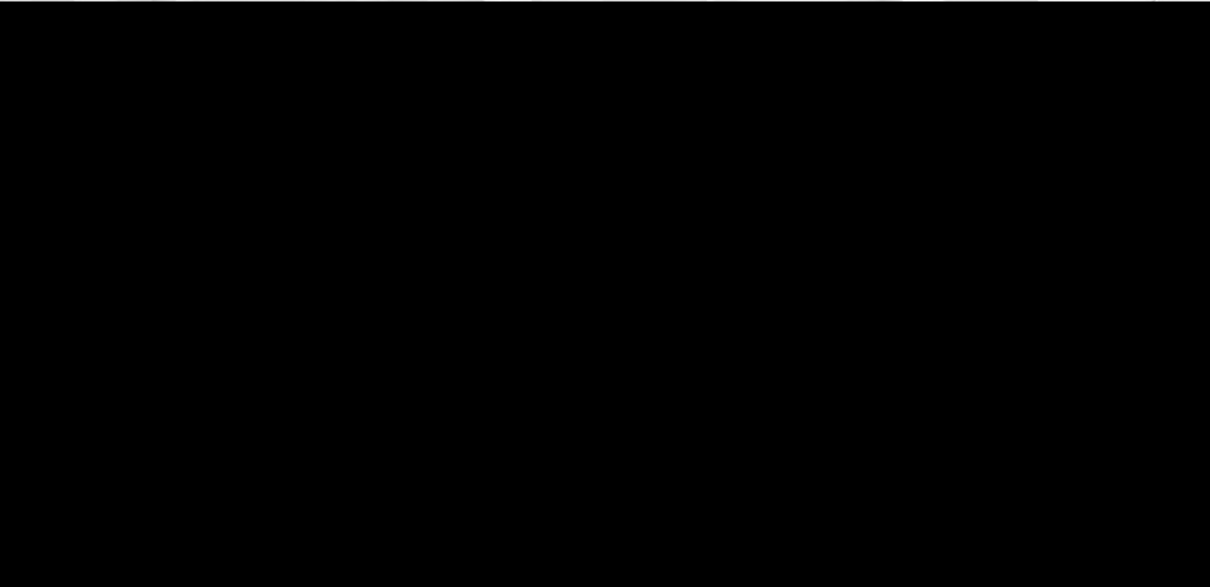
Example: Face recognition

two faces with arrows

Example: Face recognition

overview figure of the SIAM paper

Example: System identification



A brief overview of SVD

Singular value decomposition

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$

Generalization of the *eigenvalue decomposition* to **non-square matrices** by E. Beltrami (1873) and C. Jordan (1874). The first efficient numerical algorithm was developed by G. Golub *et al.* in the late 1960s.

Singular value decomposition

$$\begin{bmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_i \\ \boldsymbol{v}_i \end{bmatrix} = \sigma_i \begin{bmatrix} \boldsymbol{u}_i \\ \boldsymbol{v}_i \end{bmatrix}$$

Generalization of the *eigenvalue decomposition* to **non-square matrices** by E. Beltrami (1873) and C. Jordan (1874). The first efficient numerical algorithm was developed by G. Golub *et al.* in the late 1960s.

Schematic SVD

Geometric interpretation

Ordinary least-squares

$$\underset{\boldsymbol{x}}{\text{minimize}} \; \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

$$\begin{bmatrix} \\ \\ \quad \boldsymbol{A} \quad \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{x} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \boldsymbol{b} \\ \\ \\ \end{bmatrix}$$

Ordinary least-squares

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^{\dagger}\boldsymbol{b}$$

$$\begin{bmatrix} \\ \\ & \boldsymbol{A} & \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{x} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \boldsymbol{b} \\ \\ \\ \end{bmatrix}$$

Ordinary least-squares

$$\hat{\boldsymbol{x}} = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T\boldsymbol{b}$$

$$\begin{bmatrix} \\ \\ & & \boldsymbol{A} & & \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{x} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \boldsymbol{b} \\ \\ \\ \end{bmatrix}$$

Low-rank approximation

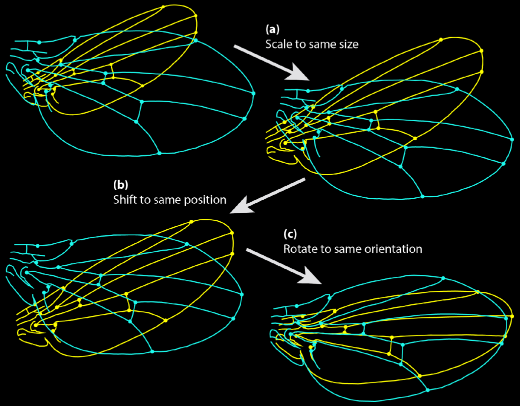
$$\begin{aligned} \underset{\hat{\boldsymbol{A}}}{\text{minimize}} \ & \|\boldsymbol{A} - \hat{\boldsymbol{A}}\|_F^2 \\ \text{subject to } & \operatorname{rank} \hat{\boldsymbol{A}} = r \end{aligned}$$

Low-rank approximation

$$\hat{\boldsymbol{A}} = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^T$$

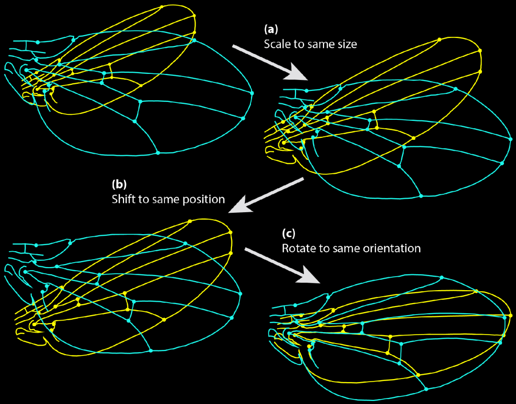
Orthogonal Procrustes problem

$$
\begin{aligned}
&\underset{\boldsymbol{R}}{\text{minimize}}\ \|\boldsymbol{R}\boldsymbol{A}-\boldsymbol{B}\|_F \\
&\text{subject to } \boldsymbol{R}^T\boldsymbol{R}=\boldsymbol{I}
\end{aligned}
$$



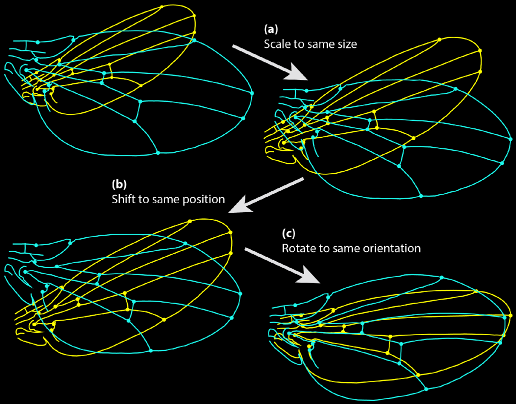
Orthogonal Procrustes problem

$$BA^T = U\Sigma V^T$$



Orthogonal Procrustes problem

$$R = UV^T$$



SVD in data-driven modeling

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{x} + \varepsilon$$

Linear deconvolution problem

Schematic of a linear system

Linear deconvolution problem

Schematic of deconvolution

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{x} + \varepsilon$$

Fundamental problem of component analysis

$$\begin{aligned} &\underset{\boldsymbol{K}}{\text{minimize}} \, \|\boldsymbol{M}^{\frac{1}{2}} \left(\boldsymbol{Y} - \boldsymbol{K}\boldsymbol{X}\right)\|_F \\ &\text{subject to rank } \boldsymbol{K} = r. \end{aligned}$$

Fundamental problem of component analysis

$$\begin{aligned}\underset{\boldsymbol{P},\boldsymbol{Q}}{\text{minimize}}\ & \|\boldsymbol{M}^{\frac{1}{2}}\left(\boldsymbol{Y}-\boldsymbol{P}\boldsymbol{Q}^T\boldsymbol{X}\right)\|_F \\ \text{subject to}\ & \boldsymbol{P}^T\boldsymbol{M}\boldsymbol{P}=\boldsymbol{I}_r \quad \left(\text{ou}\ \boldsymbol{Q}^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{Q}=\boldsymbol{I}_r\right).\end{aligned}$$

$$\begin{bmatrix} \boldsymbol{M} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_i \\ \boldsymbol{q}_i \end{bmatrix} \lambda_i = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{M}\boldsymbol{C}_{\boldsymbol{yx}} \\ \boldsymbol{C}_{\boldsymbol{xy}}\boldsymbol{M} & -\boldsymbol{C}_{\boldsymbol{xx}} \end{bmatrix} \begin{bmatrix} \boldsymbol{p}_i \\ \boldsymbol{q}_i \end{bmatrix}$$

Proper Orthogonal Decomposition

POD eigenproblem

$$M p_i \lambda_i = C_{xx} p_i$$

POD is recovered if $X = Y$ and $P = Q$. It aims at characterizing the second order statistics of the random variable x.

Complex Ginzburg-Landau equation

$$\frac{\partial q}{\partial t} = -U\frac{\partial q}{\partial x} + \mu(x)q + \gamma\frac{\partial^2 q}{\partial x^2} - \beta|q|^2 q$$

Covariance matrix

Eigenspectrum

POD mode

Phase portrait

Galerkin projection