

# M2 MMS : Réseaux de neurones pour la modélisation

7 novembre 2025

## Table des matières

<b>1</b>	<b>Machines et réseaux de neurones</b>	<b>2</b>
1.1	Machines et exemples classiques . . . . .	2
1.2	MLP ('multilayer perceptron') . . . . .	3
<b>2</b>	<b>Pourquoi construire des machines ?</b>	<b>6</b>
2.1	Différentes classes de problèmes . . . . .	6
<b>3</b>	<b>Fonction de perte</b>	<b>8</b>
3.1	Classification binaire . . . . .	8
3.2	Classification . . . . .	8
3.3	Moindre carré . . . . .	9
3.4	Remarques . . . . .	9
<b>4</b>	<b>Apprentissage</b>	<b>9</b>
4.1	Méthode de gradient . . . . .	9
4.2	Calcul différentiel . . . . .	10
4.3	Différentiation automatique . . . . .	11
4.4	Apprentissage par étape . . . . .	11
<b>5</b>	<b>EDO</b>	<b>11</b>
5.1	Méthodes de collocation . . . . .	11
5.2	Méthode variationnelle . . . . .	12
5.3	Méthode de Galerkin . . . . .	13
5.4	Fonction loss PINN . . . . .	13
<b>6</b>	<b>Réseaux pour les EDO</b>	<b>13</b>
<b>7</b>	<b>EDO avec paramètres</b>	<b>14</b>
<b>A</b>	<b>Preuves de section 1 (Machines et réseaux de neurones)</b>	<b>15</b>

---

<b>B Preuves de section 3 (Fonction de perte)</b>	<b>16</b>
<b>C Preuves de section 4 (Apprentissage)</b>	<b>16</b>
<b>D Preuves de section 5 (EDO)</b>	<b>17</b>

## Introduction

### Objectifs du cours

- Connaître les bases des machines et réseaux de neurones
- Savoir les utiliser dans le contexte de modèles basées sur les équations différentielles ordinaires (EDO)
- Savoir utiliser des librairies de l'IA en python.

Chaque chapitre correspond à 1-3 cours/TP.

### Librairies

- tensorflow
- pytorch
- jax (avec Flax et Optax)

### Prérequis

- Les base de python
- list, dict, class
- numpy

### Notations

Pour des vecteurs  $x, y \in \mathbb{R}^n$  on utilise  $x^\top y = \sum_{i=1}^n x_i y_i$  pour le produit scalaire euclidien est  $|x| = \sqrt{x^\top x}$  pour la norme associée.

Pour  $f : \mathbb{R} \rightarrow \mathbb{R}$  on définit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  (sans distinction de notation!) composante par composante,  $(f(x))_i = f(x_i), 1 \leq i \leq n$ .

## 1 Machines et réseaux de neurones

### 1.1 Machines et exemples classiques

**Définition 1.1.** Soit  $m, n, n_W \in \mathbb{N}$ . Une **machine** est une application

$$\Phi : \mathbb{R}^n \times \mathbb{R}^{n_W} \rightarrow \mathbb{R}^m, \quad (x, \tilde{w}) \rightarrow y = \Phi(x, \tilde{w}).$$

**L'apprentissage** consiste à déterminer  $\tilde{w} = w_*$  de sorte que l'application  $x \mapsto \Phi(x, w_*)$  permet de représenter des données ou un modèle physique.

On appelle  $\Phi$  une **machine vectoriel**, si

$$\Phi(x, \tilde{w}) = \sum_{i=1}^N c_i \phi_i(x, w).$$

Dans ce cas nous avons  $\tilde{w} = (c, w)$  et  $X_\Phi := \text{Im } \Phi = \text{vect } \{\phi_i(x, w) \mid 1 \leq i \leq N\}$ . Les fonction  $\phi_i^w : x \mapsto \phi_i(x, w)$  peuvent former une base ou non. On les appelle souvent 'features' en anglais. Nous avons alors une application ('feature map')

$$\mathbb{R}^{n_w} \rightarrow \{\mathbb{R}^n \rightarrow \mathbb{R}^m\}^N \quad w \mapsto (\phi_i^w)_{1 \leq i \leq N} \quad (1)$$

**Exemple 1.2.** Voici quelques exemples de machines vectorielles.

- L'interpolation de Lagrange d'une fonction univariée continue  $f : [0; 1] \rightarrow \mathbb{R}$  avec les points d'interpolation  $0 = t_0 < \dots < t_{i-1} < t_i < t_N = 1$ .

$$I_n f(t) = \sum_{i=0}^N c_i \phi_i(t), \quad \phi_i(t) = \prod_{\substack{j=1 \\ j \neq i}}^N \frac{t - t_j}{t_i - t_j}.$$

Ici, nous avons " $w = \emptyset$ ".

- Même chose, mais avec les points d'interpolation variables, donc " $w = (t_i)$ ".
- Espace  $P^1([0; 1])$ . Similaire au deux précédents. On a une base locale, c'est à dire le support de chaque  $\phi_i$  est autour de  $t_i$ .
- Espace des séries de Fourier tronquées. Similaire aux précédents. Bases globales, le support de chaque  $\phi_i$  est tout le domaine.

## 1.2 MLP ('multilayer perceptron')

**Définition 1.3.** On appelle **fonction d'activation** une fonction  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ .

1. *Sigmoide*  $\sigma(x) = \frac{1}{1+e^{-x}}$ ,  $\sigma(\mathbb{R}) \subset ]0; 1[$ .
2. *Tanh*  $\sigma(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ ,  $\sigma(\mathbb{R}) \subset ]-1; 1[$ .
3. *ReLU*  $\text{relu}(x) = x^+ = \max \{0, x\}$ ,  $\sigma(\mathbb{R}) \subset ]-1; 1[$ .
4. *Swish*  $\text{swish}(x) = x\sigma(\beta x)$ ,  $\sigma(\mathbb{R}) \subset ]-1; 1[$ .
5. *Id*  $\sigma(x) = x$ ,  $\sigma(\mathbb{R}) = \mathbb{R}$ .

Plein d'autres sont possibles. Attention : l'image de  $\sigma$  est importante.

**Définition 1.4.** Un réseau à une couche ('perceptron')  $(W, b, \sigma)$ ,  $W \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  est l'application

$$\Phi(x, W, b, \sigma) : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad x \mapsto \sigma(Wx + b). \quad (2)$$

Un réseau multi-couches (MLP, 'multi-layer-perceptron') est la composition de réseau à une couche  $\Phi_i = \Phi_i(W_i, b_i, \sigma_i)$ ,  $1 \leq i \leq L$  avec

$$\Phi(x, \tilde{w}) : \mathbb{R}^n \rightarrow \mathbb{R}^m \quad \Phi(x, \tilde{w}) = \Phi_L \circ \cdots \circ \Phi_1. \quad (3)$$

Évidemment nous avons  $\tilde{w} = (W_i, b_i)_{1 \leq i \leq L}$  et

$$W_i \in \mathbb{R}^{n_i \times n_{i-1}}, \quad b_i \in \mathbb{R}^{n_i}, \quad n_0 = n, \quad n_L = m.$$

**Remarque 1.5.** Un MLP est une machine vectoriel, si  $\sigma_L = \text{id}$ . Nous avons  $N = n_{L-1}$ , si  $b_L = 0$  et  $N = n_L$  sinon. Le fait d'avoir un biais  $b_L \neq 0$  rajoute la fonction de base  $\phi_0 = 1$ , donc, dans ce cas  $b_L$  joue aussi le rôle d'un coefficient.

Voici quelques types de couches.

- 'Linear layer' avec poids ('weight')  $W$  et biais ('bias')  $b$ .  $x \mapsto Wx + b$ .
- 'Activation layer'.  $x \mapsto \sigma(x)$ .
- 'Convolutional layer'. Une convolution s'écrit comme une multiplication matrice-vecteur :

$$(1D) \quad y_i = \sum_{k=-m}^m w_k x_{i+k}, \quad (2D) \quad y(i, j) = \sum_{k=-m}^m \sum_{l=-m}^m w(k, l) x(i+k, j+l).$$

- 'Residual layer'  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\Phi(x) = x + \sigma(Wx + b)$
- 'Attention layer' :  $\Phi(x) = \text{softmax}(n^{-1/2} K Q^\top) V$  ( $K, Q, V \in \mathbb{R}^{n \times n}$ )

$$\text{softmax}_\beta : \mathbb{R}^n \rightarrow ]0; 1[^n \quad x \mapsto \left( \frac{e^{\beta x_i}}{\sum_{j=1}^n e^{\beta x_j}} \right)_{1 \leq i \leq n}.$$

C'est le gradient de logsumexp $_\beta : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $x \mapsto \beta \log \left( \sum_{j=1}^n e^{\beta x_j} \right)$

- 'Dropout layer' Il s'agit de réduire le nombre de paramètres de la matrice  $W$ .
- 'Implicit layer' Ici  $\Phi$  correspond à la solution d'une équation, par exemple une équation différentielle.
- 'SVM' exemple d'un problème d'optimisation, aussi 'implicit'.

**Exercice 1.1.** Soit  $a < b$ . Une partition  $\mathcal{T}$  de l'intervalle est un ensemble ordonné de  $r \in \mathbb{N}_{\geq 1}$  points  $x_i$  tel que

$$\mathcal{T} : a = x_0 < x_1 < \dots < x_n = b, \quad I_k = [x_{k-1}, x_k[$$

Une fonction  $\varphi : [a; b] \rightarrow \mathbb{R}$  est affine par morceaux sur la partition  $\mathcal{T}$  de  $[a; b]$  s'il existe des nombres réels  $\alpha_k, \beta_k$  tel que pour  $k = 1, \dots, n$  nous avons

$$\varphi(x) = \alpha_k + \beta_k x \quad \text{pour } x \in I_k, \quad \varphi(b) = \alpha_n + \beta_n b. \quad (4)$$

Soit  $\mathcal{T}$  une partition de  $[a; b]$ . On définit

$$\mathcal{P}_{\mathcal{T}}^1 := \{\varphi \in C([a; b]) \mid \varphi \text{ est affine par morceaux sur } \mathcal{T}\}. \quad (5)$$

Montrer que

1.  $\mathcal{P}_{\mathcal{T}}^1$  est un espace vectoriel.
2. Soit  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  écrite sous la forme (4). Quelles conditions doivent satisfaire les paramètres  $\alpha_k$  et  $\beta_k$  ?
3. Il existe des fonctions  $\varphi_k \in \mathcal{P}_{\mathcal{T}}^1$ ,  $0 \leq k \leq n$  telles que

$$\varphi_k(x_\ell) = \begin{cases} 1 & k = \ell, \\ 0 & k \neq \ell. \end{cases} = \delta_{k\ell} \quad (\text{symbole de Kronecker}). \quad (6)$$

4. Toute fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  s'écrit comme

$$\varphi(x) = \sum_{k=0}^{n+1} c_k \varphi_k(x) \quad c_k \in \mathbb{R}. \quad (7)$$

En particuliers, déterminer les  $c_k$  pour  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  de la forme (4).

5. Toute fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  s'écrit comme

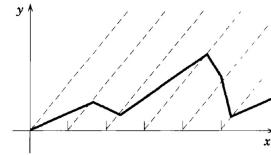
$$\varphi(x) = \alpha + \beta x + \sum_{k=1}^{n-1} d_k (x - x_k)_+ \quad (8)$$

avec  $\alpha, \beta, d_k \in \mathbb{R}$ .

6. Toute fonction  $f : [a; b] \rightarrow \mathbb{R}$  continue peut être approchée de façon uniforme par  $\mathcal{P}_{\mathcal{T}}^1$ , c'est à dire qu'il existe pour tout  $\varepsilon > 0$  une fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  de sorte que

$$|f(x) - \varphi(x)| \leq \varepsilon \quad \forall a \leq x \leq b.$$

7. Écrire un MLP qui permet de construire l'espace  $P^1(h)$  sur un maillage 1D  $h = (a = x_0, \dots, x_n = b)$ .



8. Se servir de la formule pour  $k \geq 1$

$$\sum_{j=0}^k b_j x^j + \sum_{i=1}^{n-1} a_i (x - x_i)_+^k \quad (9)$$

pour construire un MLP pour l'espace des B-splines  $\mathcal{B}_T^k$ .

9. Comment construire un MLP pour le  $P^1(h)$  sur un maillage en triangles ?

## 2 Pourquoi construire des machines ?

- Expliquer les relations entre  $x$  et  $y$  : dans le cas unidimensionnel, on interprète de façon géométrique et algébrique  $y = ax$  ou  $y = bx^2$ .
- Prédire le résultat  $y$  associé à un  $x$ .

### 2.1 Différentes classes de problèmes

- **Classification (classement)**

- ▷ binaire : il y a seulement deux caractéristiques (que l'on peut appeler 0 et 1,  $y \in \{0, 1\}$  ou -1 et +1) : chat ou chien, appareil fonctionne ou pas, passager survit à la catastrophe du *Titanic*,...
- ▷ classes multiples ('pattern recognition') : reconnaître un chiffre donné  $y \in \{0, 1, \dots, 9\}$ , différents types de vélos,...

- **Régression** :  $y \in \mathbb{R}$  dans le cas scalaire. Étant donné  $(x_i, y_i)_{1 \leq i \leq n_D}$  on cherche  $f$  tel que

$$l(f) := \frac{1}{2n_D} \sum_{i=1}^{n_D} (y_i - f(x_i))^2$$

soit le plus petit possible. Ce problème dépend de la classe de fonction  $f \in \mathcal{F}$ . Si  $\mathcal{F}$  est assez grand nous rappelons que  $l(f_{\text{lag}}) = 0$  pour le polynôme d'interpolation de Lagrange  $f_{\text{lag}}$ , car interpolation signifie  $y_i = f(x_i)$ . On se rappelle également que ce processus est instable pour  $n_D \rightarrow +\infty$  (et même  $n_D \approx 10$ , phénomène de Runge). On s'intéresse plutôt à une classe  $\mathcal{F}$  avec peu de paramètre. Par exemple Gauß avait pris  $\mathcal{F} = \{x \mapsto a + bx \mid a, b \in \mathbb{R}\}$ , ce qui permet une interprétation des paramètres.

- **Simplification de modèle** : étant donné un modèle compliqué  $\Phi^*$  que l'on ne peut évaluer, on cherche une approximation  $\Phi$ . Pour l'exemple d'une EDO nonlinéaire (20)

$$\Phi^* : U \subset \mathbb{R}^n \rightarrow C^1(I, \mathbb{R}^n), \quad \Phi^*(u_0) = \text{solution de (20) sur l'intervalle } I$$

Premièrement,  $\Phi^*$  ne prend pas des valeurs en un espace de dimension finie. Deuxièmement, on ne connaît pas  $\Phi^*(u_0)(t)$  pour un temps  $t \in I$ . On est obligé de calculer une approximation par une méthode numérique. Et on peut utiliser un réseaux de neurones, qui devient alors une méthode numérique. Il y a différentes techniques qui rentre dans cette catégorie

- ▷ **Méthodes de Galerkine** (Séries de Fourier, FEM, ondelettes, méthodes spectrales) où l'on construit des fonctions de base appropriée pour le modèle en question. Ces méthodes dépendent de paramètres. Dans le cas des séries de Fourier il s'agit des indices que l'on utilise et dans le cas de la FEM, il s'agit par exemple du maillage. Depuis longtemps il existe des algorithmes qui cherche à optimiser ces paramètres (méthodes adaptatives) en fonction de la 'solution'  $\Phi^*(u_0)$ . Il s'agit de l'approximation non-linéaire.  
Une caractéristique des méthode de Galerkine est que les coefficients sont obtenus par la résolution d'un système d'équations nonlinéaires. Dans les réseaux de neurones, même si la machine est vectorielle, on minimise une fonction perte.
- ▷ **Méthode de réduction de modèle**, POD où on essaie d'extraire l'information essentielle pour construire une base adaptée. Ces méthodes sont souvent combinées avec une méthode de Galerkine.

- **Mélange de régression et de réduction de modèle.** Si nous avons un modèle  $\Phi^*$  et des observations (données), on peut facilement combiner les deux si le problème d'apprentissage est une minimisation : il suffit de rajouter une mesure des écarts entre modèle et observations à la fonction à minimiser (la perte). Cela est souvent utilisé comme argument en faveur des réseaux de neurones. Toutefois, minimiser la somme de deux fonctions  $f$  et  $g$  dépend d'une pondération que l'on choisit a priori et le résultat peut être très différent des résultats des autres problèmes qui se posent naturellement

$$\begin{cases} \min \{f(x) + \alpha g(x) \mid x \in \mathbb{R}^n\}, \\ \min \{f(x) \mid x \in \mathbb{R}^n \cap \{\|g(x)\| \leq \varepsilon\}\}, \\ \min \{g(x) \mid x \in \mathbb{R}^n \cap \{\|f(x)\| \leq \delta\}\}. \end{cases}$$

### 3 Fonction de perte

**Définition 3.1.** On appelle **données** (*data*) un ensemble  $\mathcal{D} = (x_i, y_i)_{1 \leq i \leq n_D}$ ,  $x_i \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^m$ . Alors  $\mathcal{D} \subset \mathbb{R}^{n_D} \approx (\mathbb{R}^n \times \mathbb{R}^m)^{n_D}$ . Souvent, on partage les données en données d'apprentissage et données de test (et éventuellement données de validation).

**Définition 3.2.** On appelle **fonction de perte** (*loss*) une fonction  $\tilde{L} : \mathbb{R}^{n_W} \times \mathbb{R}^{n_D} \rightarrow \mathbb{R}$  dont la minimisation est l'**apprentissage** (*learning*). On suppose avoir  $L : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$  telle que

$$\tilde{L}(w, \mathcal{D}) = \frac{1}{n_D} \sum_{i=1}^{n_D} L(\Phi(x_i, w), y_i). \quad (10)$$

Le problème d'apprentissage est alors la minimisation de  $\hat{L} : \mathbb{R}^{n_W} \rightarrow \mathbb{R}$

$$\min \left\{ \hat{L}(w) \mid w \in \mathbb{R}^p \right\}, \quad \hat{L}(w) = \tilde{L}(w, \mathcal{D}). \quad (11)$$

**Remarque 3.3.** On peut avoir une régularisation  $\hat{L}(w) = \tilde{L}(w, \mathcal{D}) + \frac{\alpha}{2} \|w - w_0\|^2$  avec  $w_0$  des valeurs données des paramètres du modèle, ou, encore plus complexe, des contraintes sur les paramètres du modèle.

**Exemple 3.4.** • (moindre carrés,  $l^2$ )  $L(z, y) = \frac{1}{2} \|y - z\|^2$ .  
• ( $l^1$ )  $L(z, y) = \|y - z\|_{l^1(\mathbb{R}^m)}$ .

#### 3.1 Classification binaire

On utilise souvent la 'binary cross entropy' (si la dernière couche finit avec une sigmoïde)

$$y_i \in \{-1, +1\}, \quad l(x) := -\frac{1}{n_D} \sum_{i=1}^{n_D} y_i \log(z_i) + (1-y_i) \log(1-z_i); \quad z_i = \Phi(x_i, w) \quad (12)$$

#### 3.2 Classification

Avec  $n_C$  classes,  $y_{ij} \geq 0$ ,  $\sum_{j=1}^{n_C} y_{ij} = 1$  pour tout  $i$ .

$$y_i \in [\![1, \dots, n_c]\!] \quad l(x) := -\frac{1}{n_D} \sum_{i=1}^{n_D} \sum_{j=1}^{n_C} y_{ij} \log(z_{ij}); \quad z_{ij} = \Phi(x_i, w). \quad (13)$$

On interprète les  $z_{ij}$  comme la probabilité que

### 3.3 Moindre carré

$$l(x) = \frac{1}{2} \sum_{i=1}^m |y_i - \Phi(x_i, w)|^2. \quad (14)$$

### 3.4 Remarques

Il y a une certaine ambiguïté dans nos définition de couches et de fonction de perte : on peut toujours intégrer les dernières couches dans la fonction de perte.

**Théorème 3.5.** *Si la machine contient une couche de la forme  $x \rightarrow \sigma(Wx + b)$ , on peut modifier la fonction de perte telle que la machine est vectorielle.*

**Exemple 3.6.** Pour les problèmes de classification on utilise généralement la fonction logistique.

## 4 Apprentissage

**Définition 4.1.** Un problème d'apprentissage de machine 'machine learning' consiste en trois ingrédients

- Une machine  $\Phi(x, w)$
- Une fonction de perte (ce qui implique généralement les données et pratiquement leur traitement)
- Un algorithme d'optimisation (y compris critère d'arrêt).

Ces ingrédients étant fixés, l'apprentissage consiste alors à minimiser la fonction, voir (10) et (11)

$$f : \mathbb{R}^{n_w} \rightarrow \mathbb{R}, \quad w \mapsto f(w) := \hat{L}(w) \left( = \frac{1}{n_D} \sum_{i=1}^{n_D} L(\Phi(x_i, w), y_i) \right). \quad (15)$$

Cette fonction est différentiable dans beaucoup de cas (sauf ReLU), mais d'autres propriétés comme la convexité ne sont souvent pas présente.

### 4.1 Méthode de gradient

En général, les méthodes numériques utilisées dans l'apprentissage sont de variation de la méthode du gradient (ou de la descente). Partant d'un ensemble de paramètres  $w_0 \in \mathbb{R}^{n_w}$ , on définit

$$w_{k+1} := w_k + t_k \nabla f(w_k), \quad t_k \text{ est le pas ou 'learning rate'} \quad (16)$$

Une itération est souvent appelé 'epoch'. L'ingrédient essentiel dans (16) est la gradient.

**Définition 4.2.** Soit  $(X, \langle \cdot, \cdot \rangle)$  un espace de Hilbert et  $f \in C^1(X, \mathbb{R})$ . Alors

$$\langle \nabla f(x), p \rangle := \lim_{s \searrow 0} \frac{f(x + sp) - f(x)}{s}. \quad (17)$$

On note que  $\nabla f(x) = Df(x)^\top$  (voir définition 4.3).

On voit qu'une difficulté majeure est le calcul de gradient.

## 4.2 Calcul différentiel

On rappelle les notations du calcul différentiel.

**Définition 4.3.**  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  est différentiable en  $x$  s'il existe  $A \in \mathbb{R}^{m \times n}$  tel que

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\|R(x, h)\|}{\|h\|} = 0 \quad (R(x, h) := F(x + h) - F(x) - Ah), \quad DF(x) := A.$$

Pour des fonctions de plusieurs variables, par exemple  $F(x, z)$  on écrit  $DF_x(x, z)$  et  $DF_z(x, z)$  pour les dérivées "partielles".

On rappelle également

**Théorème 4.4.** Soit  $m, n, p \in \mathbb{N}$ ,  $U \subset \mathbb{R}^n$ ,  $V \subset \mathbb{R}^m$ ,  $F \in C^1(U, \mathbb{R}^m)$ ,  $G \in C^1(V, \mathbb{R}^p)$  et  $F(U) \subset V$ . Alors  $h := g \circ f \in C^1(U, \mathbb{R}^p)$  et pour  $x \in U$  et  $y := f(x) \in V$

$$Dh(x) = Dg(y)Df(x), \quad \begin{array}{ccc} U & \xrightarrow{f} & V \\ & \searrow h & \downarrow g \\ & & \mathbb{R}^m \end{array}, \quad \begin{array}{ccc} \mathbb{R}^n & \xrightarrow{Df(x)} & \mathbb{R}^n \\ & \searrow Dh(x) & \downarrow Dg(y) \\ & & \mathbb{R}^m \end{array}. \quad (18)$$

Par les théorèmes du calcul différentiel nous avons.

**Théorème 4.5.** Si  $L$  et  $\Phi$  sont continûment différentiables, alors  $\tilde{L} \in C^1(\mathbb{R}^{n_w})$ , le gradient de  $\tilde{L}$  s'écrit

$$\nabla \tilde{L}(w) = \frac{1}{n_D} \sum_{i=1}^{n_D} \Phi'_w(x_i, w)^\top \nabla_z L(\Phi(x_i, w), y_i) \quad (19)$$

### 4.3 Différentiation automatique

### 4.4 Apprentissage par étape

## 5 EDO

On considère  $U \subset \mathbb{R}^n$  ouvert,  $u_0 \in U$  et une fonction  $f \in C^1(U, \mathbb{R}^n)$ . L'équation différentielle ordinaire (EDO) d'ordre un

$$\begin{cases} \frac{du}{dt} = f(t, u) \\ u(0) = u_0. \end{cases} \quad (20)$$

On rappelle le théorème de Cauchy-Lipschitz. S'il existe  $r > 0$  et  $L > 0$  tel que

$$|f(u) - f(v)| \leq L |u - v| \quad \forall u, v \in U \cap B_r(u_0), \quad (21)$$

alors il existe  $T > 0$  et une solution unique de (20) sur  $[0, T]$ .  $f \in C^1(U, \mathbb{R}^n)$  est suffisant pour (21).

**Remarque 5.1.** 1. Il existe un temps maximal  $T \in \mathbb{R} \cup \{+\infty\}$  d'existence de solution. Si  $T < +\infty$ , nous avons  $\lim_{t \nearrow T} |u(t)| = +\infty$ .

2. L'équation (20) s'écrit comme une EDO autonome avec une variable supplémentaire

$$\frac{dv}{dt} = g(v), \quad v(0) = v_0$$

avec  $v = (u, t)$ ,  $g(v) = (f(u), 1)$  et  $v_0 = (u_0, 0)$ , la variable supplémentaire est le temps.

### 5.1 Méthodes de collocation

Une méthode de collocation sur un intervalle  $I = [0; T]$  consiste à choisir un espace  $X_N \subset C^1(I, \mathbb{R}^n)$  de dimension  $N + 1$  et des points (maillage)

$$\tau = (0 = t_0 < t_1 < \dots < t_N = T). \quad (22)$$

On impose les équations à une fonction  $u_N \in X_N$

$$u(0) = u_0, \quad \frac{du}{dt}(t_i) = f(u(t_i)) \quad 1 \leq i \leq N. \quad (23)$$

Soit  $(\phi_j)_{0 \leq j \leq N}$  une base de  $X_N$ . (23) est alors converti en un système algébrique pour les coefficient  $c \in \mathbb{R}^{N+1}$

$$u(t) = \sum_{j=0}^N c_j \phi_j(t), \quad \sum_{j=0}^N c_j \phi_j(0) = u_0, \quad \sum_{j=0}^N c_j \frac{d\phi_j}{dt}(t_i) = f\left(\sum_{j=0}^N c_j \phi_j(t)\right).$$

**Remarque 5.2.** Pour une équation linéaire  $f(u) = Au + b$  on obtient un système linéaire à résoudre

$$Kc = \gamma, \quad K_{ij} = \frac{d\phi_j}{dt}(t_i) - A\phi_j(t_i), \quad \gamma_i = b \quad (24)$$

**Exercice 5.1.** On rappelle (voir le cours d'analyse numérique M1MMS) que

- i) pour deux nombres réels  $a < b$  un **maillage**  $h = (t_i)_{0 \leq i \leq N}$  est une partition de l'intervalle  $[a; b]$

$$a = t_0 < t_1 < \cdots < t_n = b, \quad I_j := [t_{j-1}; t_j[, \quad (1 \leq j \leq N),$$

- ii) l'espace des **B-splines** d'ordre  $k \in \mathbb{N}^*$  sur un maillage  $h = (t_i)_{0 \leq i \leq N}$  est défini par

$$\mathcal{B}_h^k(I) := \left\{ v \in C^{k-1}([a; b]) \mid v|_{I_j} \in P^k(I_j) \quad 1 \leq j \leq N \right\}. \quad (25)$$

1. Pour  $X_N = \mathcal{B}_h^2(I)$ , l'espace des splines quadratiques de classe  $C^1$  sur le **même** maillage  $h = (t_i)_{0 \leq i \leq N}$  on obtient un schéma de type Crank-Nicolson pour une EDO linéaire à coefficients constants et  $h$  uniforme.
2. Faire une analyse de l'erreur de la méthode.

**Remarque 5.3.** Les méthodes de collocation font partie des méthodes de Runge-Kutta implicites.

## 5.2 Méthode variationnelle

Le théorème de Cauchy-Lipschitz cherche une solution dans l'espace de Banach  $C([0; T], \mathbb{R}^n)$ . Ensuite cette solution sera même  $C^1([0; T], \mathbb{R}^n)$  si  $f$  est continue.

Ici on va chercher une solution dans  $X := H^1([0; T], \mathbb{R}^n)$  telle que pour tout  $(v, v_0) \in Y := L^2([0; T], \mathbb{R}^n) \times \mathbb{R}^n$

$$u \in X : \int_0^T v(s)^\top \frac{du}{dt}(s) ds + v_0^\top u(0) = \int_0^T v(s)^\top f(s, u(s)) ds + v_0^\top u(0). \quad (26)$$

On peut baser la démonstration de Cauchy-Lipschitz sur (26), mais ce n'est pas l'objectif ici. Par contre, voyons le cas affine-autonome,  $f(t, u) = b - Au$ . Dans ce cas l'existence (et l'unicité) de la formulation (26) est assuré si la forme bilinéaire

$$a(u, (v, v_0)) := \int_0^T v(s)^\top \left( \frac{du}{dt}(s) + Au(s) \right) ds + v_0^\top u(0)$$

satisfait la condition inf-sup (voir le cours d'analyse numérique M2MMS)

$$\inf_{u \in B_X} \sup_{(v, v_0) \in B_Y} a(u, (v, v_0)) = \inf_{(v, v_0) \in B_Y} \sup_{u \in B_X} a(u, (v, v_0)) = \gamma > 0. \quad (27)$$

### 5.3 Méthode de Galerkin

Une méthode de Galerkin (conforme) consiste à définir deux espaces  $X_N \subset$  et  $Y_N \subset Y$  et chercher  $u \in X_N$  tel que

$$\int_0^T v(s)^\top \frac{du_N}{dt}(s) ds + v_0^\top u_N(0) = \int_0^T v(s)^\top f(s, u_N(s)) ds + v_0^\top u(0) \quad (28)$$

pour tout  $(v, v_0) \in Y_N$ .

### 5.4 Fonction loss PINN

Une première idée est d'utiliser un espace  $\tilde{X} = X_\Phi$  généré par une machine vectorielles  $\Phi$  est d'utiliser la fonction perte

$$\frac{1}{2} \|\tilde{u}(0) - u_0\|^2 + \sum_{i=0}^N \frac{1}{2} \left\| \frac{d\tilde{u}}{dt}(t_i) - f(\tilde{u}(t_i)) \right\|^2$$

Il est alors facile de rajouter un term avec des données  $\mathcal{D} = (\tilde{t}_k, \tilde{w}_k)$  (mesures expérimentales)

$$l_{\text{PINN}}(\tilde{u}) = \frac{1}{2} \|\tilde{u}(0) - u_0\|^2 + \sum_{i=0}^N \frac{1}{2} \left\| \frac{d\tilde{u}}{dt}(t_i) - f(\tilde{u}(t_i)) \right\|^2 + \sum_{k=1}^d \frac{1}{2} \|\tilde{u}(\tilde{t}_k) - \tilde{w}_k\|^2$$

**Remarque 5.4.** Dans le cas sans données ( $d = 0$ ), si le MLP produit  $B^1(I)$ , la machine produit la solution de Crank-Nicolson. Une difficulté est le conditionnement du problème.

## 6 Réseaux pour les EDO

1. Pour l'exemple de l'EDO d'ordre deux scalaire
  - (a) Que se passe-t-il si l'on a plus de fonctions de base que de points de collocations ?
  - (b) Étudier la convergence et l'erreur en fonction du nombre de point de collocation (et du nombre de fonctions de base).
  - (c) Changer les conditions de limites : on remplace  $u(T) = 0$  par  $u'(0) = \alpha$ .
  - (d) Que se passe-t-il si l'on varie le nombre de couches ?
  - (e) Que se passe-t-il si l'on change la fonction d'activation ?

- (f) Que se passe-t-il si l'on change la perte : par exemple pondération différente,  $l_1$  ?
- (g) Essayer d'implémenter les conditions de bord comme contraintes.
- (h) Plotter les fonctions de base

## 7 EDO avec paramètres

Dans la pratique, les modèles mathématiques contiennent des paramètres (physique ou non). Soit  $n_p \in \mathbb{N}$  et  $p = (p_0, p_1) \in \mathbb{R}^{n_p}$  et

$$\begin{cases} \frac{du}{dt} = f(u, p_1) \\ u(0) = u_0(p_0). \end{cases} \quad (29)$$

Nous avons alors une application

$$S : \mathbb{R}^{n_p} \rightarrow C^1(I, \mathbb{R}^n) \quad p \rightarrow u(p). \quad (30)$$

Des questions typiques sont

- Déterminer des paramètres à partir de mesures.
- Plus modestement : déterminer la sensibilité des solution par rapport aux paramètre.
- Plus ambitieux : déterminer les mesures les plus importantes.
- Dans un autre registre : trouver des valeurs critiques des paramètre. Les valeurs critiques sont celles quand la solution  $p \rightarrow u(p)$  change de comportement, par exemple des points de bifurcation.

La clé à toutes ces questions est l'étude de l'application  $S$  définie en (30). Si elle est différentiable nous avons

$$\begin{cases} u := S(p), \quad \delta u := S'(p)(\delta p) \\ \begin{cases} \frac{d\delta u}{dt} = f'_u(u, p_1)\delta u + f'_p(u, p_1)\delta p_1 \\ \delta u(0) = u'_0(p_0)(\delta p_0). \end{cases} \end{cases} \quad (31)$$

## A Preuves de section 1 (Machines et réseaux de neurones)

**Exercice 1.1.** Soit  $a < b$ . Une partition  $\mathcal{T}$  de l'intervalle est un ensemble ordonné de  $r \in \mathbb{N}_{\geq 1}$  points  $x_i$  tel que

$$\mathcal{T} : a = x_0 < x_1 < \dots < x_n = b, \quad I_k = [x_{k-1}, x_k[$$

Une fonction  $\varphi : [a; b] \rightarrow \mathbb{R}$  est affine par morceaux sur la partition  $\mathcal{T}$  de  $[a; b]$  s'il existe des nombres réels  $\alpha_k, \beta_k$  tel que pour  $k = 1, \dots, n$  nous avons

$$\varphi(x) = \alpha_k + \beta_k x \quad \text{pour } x \in I_k, \quad \varphi(b) = \alpha_n + \beta_n b. \quad (4)$$

Soit  $\mathcal{T}$  une partition de  $[a; b]$ . On définit

$$\mathcal{P}_{\mathcal{T}}^1 := \{\varphi \in C([a; b]) \mid \varphi \text{ est affine par morceaux sur } \mathcal{T}\}. \quad (5)$$

Montrer que

1.  $\mathcal{P}_{\mathcal{T}}^1$  est un espace vectoriel.
2. Soit  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  écrite sous la forme (4). Quelles conditions doivent satisfaire les paramètres  $\alpha_k$  et  $\beta_k$  ?
3. Il existe des fonctions  $\varphi_k \in \mathcal{P}_{\mathcal{T}}^1$ ,  $0 \leq k \leq n$  telles que

$$\varphi_k(x_\ell) = \begin{cases} 1 & k = \ell, \\ 0 & k \neq \ell. \end{cases} = \delta_{k\ell} \quad (\text{symbole de Kronecker}). \quad (6)$$

4. Toute fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  s'écrit comme

$$\varphi(x) = \sum_{k=0}^{n+1} c_k \varphi_k(x) \quad c_k \in \mathbb{R}. \quad (7)$$

En particuliers, déterminer les  $c_k$  pour  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  de la forme (4).

5. Toute fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  s'écrit comme

$$\varphi(x) = \alpha + \beta x + \sum_{k=1}^{n-1} d_k (x - x_k)_+ \quad (8)$$

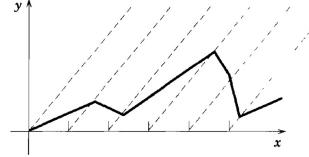
avec  $\alpha, \beta, d_k \in \mathbb{R}$ .

6. Toute fonction  $f : [a; b] \rightarrow \mathbb{R}$  continue peut être approchée de façon uniforme par  $\mathcal{P}_{\mathcal{T}}^1$ , c'est à dire qu'il existe pour tout  $\varepsilon > 0$  une fonction  $\varphi \in \mathcal{P}_{\mathcal{T}}^1$  de sorte que

$$|f(x) - \varphi(x)| \leq \varepsilon \quad \forall a \leq x \leq b.$$

7. Écrire un MLP qui permet de construire l'espace  $P^1(h)$  sur un maillage

$$1D \ h = (a = x_0, \dots, x_n = b).$$



8. Se servir de la formule pour  $k \geq 1$

$$\sum_{j=0}^k b_j x^j + \sum_{i=1}^{n-1} a_i (x - x_i)_+^k \quad (9)$$

pour construire un MLP pour l'espace des B-splines  $\mathcal{B}_T^k$ .

9. Comment construire un MLP pour le  $P^1(h)$  sur un maillage en triangles ?

Démonstration.

1. Évidemment.

2. Continuité aux  $x_i$ .

3. Voir éléments finis.

4. Voir éléments finis.

5. C'est un changement de base  $\{1, (x - x_k)_+\}$  est une autre base Nous avons

$$1 = \sum_{k=0}^{n+1} \phi_k, \quad (x - x_k)_+ = \sum_{j=k+1}^n (x_j - x_k) \phi_j(x).$$

6. Continuité uniforme.

□

## B Preuves de section 3 (Fonction de perte)

**Théorème 3.5.** Si la machine contient une couche de la forme  $x \rightarrow \sigma(Wx + b)$ , on peut modifier la fonction de perte telle que la machine est vectorielle.

Démonstration. On écrit  $x \mapsto \Phi(x, W, b) = \sigma(Wx + b)$  comme  $\Phi_2 \circ \Phi_1$  avec  $x \mapsto \Phi_1(x, W) = Wx$  et  $y \mapsto \Phi_2(y, b) = \sigma(y + b)$ . On a alors

$$L(\Phi(x, W, b), y) = L_2(\Phi_1(x, W), y), \quad L_2(x, y) := L(\Phi_2(x), y).$$

□

## C Preuves de section 4 (Apprentissage)

**Théorème 4.4.** Soit  $m, n, p \in \mathbb{N}$ ,  $U \subset \mathbb{R}^n$ ,  $V \subset \mathbb{R}^m$ ,  $F \in C^1(U, \mathbb{R}^m)$ ,  $G \in C^1(V, \mathbb{R}^p)$  et  $F(U) \subset V$ . Alors  $h := g \circ f \in C^1(U, \mathbb{R}^p)$  et pour  $x \in U$  et  $y := f(x) \in V$

$$Dh(x) = Dg(y)Df(x), \quad \begin{array}{ccc} U & \xrightarrow{f} & V \\ & \searrow h & \downarrow g \\ & & \mathbb{R}^m \end{array}, \quad \begin{array}{ccc} \mathbb{R}^n & \xrightarrow{Df(x)} & \mathbb{R}^n \\ & \searrow Dh(x) & \downarrow Dg(y) \\ & & \mathbb{R}^m \end{array}. \quad (18)$$

**Théorème 4.5.** Si  $L$  et  $\Phi$  sont continûment différentiables, alors  $\tilde{L} \in C^1(\mathbb{R}^{n_W})$ , le gradient de  $\tilde{L}$  s'écrit

$$\nabla \tilde{L}(w) = \frac{1}{n_D} \sum_{i=1}^{n_D} \Phi'_w(x_i, w)^\top \nabla_z L(\Phi(x_i, w), y_i) \quad (19)$$

## D Preuves de section 5 (EDO)

**Exercice 5.1.** On rappelle (voir le cours d'analyse numérique M1MMS) que

- i) pour deux nombres réels  $a < b$  un **maillage**  $h = (t_i)_{0 \leq i \leq N}$  est une partition de l'intervalle  $[a; b]$

$$a = t_0 < t_1 < \cdots < t_n = b, \quad I_j := [t_{j-1}; t_j[, \quad (1 \leq j \leq N),$$

- ii) l'espace des **B-splines** d'ordre  $k \in \mathbb{N}^*$  sur un maillage  $h = (t_i)_{0 \leq i \leq N}$  est défini par

$$\mathcal{B}_h^k(I) := \left\{ v \in C^{k-1}([a; b]) \mid v|_{I_j} \in P^k(I_j) \quad 1 \leq j \leq N \right\}. \quad (25)$$

1. Pour  $X_N = \mathcal{B}_h^2(I)$ , l'espace des splines quadratiques de classe  $C^1$  sur le **même** maillage  $h = (t_i)_{0 \leq i \leq N}$  on obtient un schéma de type Crank-Nicolson pour une EDO linéaire à coefficients constants et  $h$  uniforme.
2. Faire une analyse de l'erreur de la méthode.