MTH8302 - Devoir 1 : Rappel Mathématique

February 9, 2025

1. **Problème 1 : Gradients et Hessiennes** ***(30 points)***

   - **Rappel :**
     - Une matrice $\mathbf{A} \in \mathbb{R}^{n\times n}$ est symétrique si $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$.
     - Le gradient $\nabla f(x)$ d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est un vecteur contenant les dérivées partielles tel que :

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix} = \left(\frac{\partial f}{\partial \mathbf{x}}\right).$$

     - La Hessienne $\nabla^2 f(\mathbf{x})$ est une matrice symétrique $n \times n$ contenant les dérivées secondes :

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

     - **Règle de la dérivée en chaîne (matricielle) :** Pour une fonction composée $f(h(\mathbf{x}))$, la dérivée est donnée par :

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{\partial f}{\partial h} \cdot \frac{\partial h}{\partial \mathbf{x}}.$$

     - **Règle du produit matriciel :** Pour deux matrices différentiables $\mathbf{A}(\mathbf{x})$ et $\mathbf{B}(\mathbf{x})$, la dérivée est :

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{AB}) = \frac{\partial \mathbf{A}}{\partial \mathbf{x}}\mathbf{B} + \mathbf{A}\frac{\partial \mathbf{B}}{\partial \mathbf{x}}.$$

     - **Dérivée d'un produit de type vecteur-matrice-vecteur :** Si $\mathbf{b}$ et $\mathbf{x}$ sont des vecteurs et $\mathbf{A}$ est une matrice, alors la dérivée de $\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ par rapport à $\mathbf{x}$ est :

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{x}) = \mathbf{A}^{\mathrm{T}}\mathbf{b}.$$

     - **Dérivée d'une forme quadratique :** Si $\mathbf{A} \in \mathbb{R}^{n\times n}$ n'est pas nécessairement symétrique, alors la dérivée de $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ est :

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}) = (\mathbf{A} + \mathbf{A}^{\mathrm{T}})\,\mathbf{x}.$$

     - Pour justifiez votre réponse Vous pouvez suivre les étapes suivantes :

    - $*$ 1. Expansion de l'expression en termes de scalaires.
    - $*$ 2. Calcul de la dérivée de l'expression en termes de scalaires.
    - $*$ 3. Réécriture de la dérivée calculée en notation matricielle ou vectorielle.
  - Vous pouvez aussi utiliser les règles de dérivées partielles comme justification (Vous pouvez en trouver ici *The Matrix Cookbook*.)
  - Toute autre méthode de justification valide est acceptée.

- **Problème 1 : Question 1**
  - Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x} + \mathbf{b}^\mathrm{T}\mathbf{x} + c,$ où :
    - $*$ $\mathbf{A} \in \mathbb{R}^{n\times n}$ est une matrice symétrique.
    - $*$ $\mathbf{x} \in \mathbb{R}^n$ est un vecteur réel.
    - $*$ $\mathbf{b} \in \mathbb{R}^n$ est un vecteur réel.
    - $*$ $c \in \mathbb{R}$ un scalaire réel.
  - Trouvez $\nabla_\mathbf{x} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla_\mathbf{x} f(\mathbf{x}) = \frac{1}{2}\mathbf{A}\mathbf{x} + \mathbf{b}$
    (b) $\nabla_\mathbf{x} f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$
    (c) $\nabla_\mathbf{x} f(\mathbf{x}) = 2\mathbf{A}\mathbf{x} + \mathbf{b}$
    (d) $\nabla_\mathbf{x} f(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$
  - Justifiez votre réponse ***(2.5 points)***.

- **Problème 1 : Question 2**
  - Soit $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x} + \mathbf{b}^\mathrm{T}\mathbf{x} + c,$ où :
    - $*$ $\mathbf{A} \in \mathbb{R}^{n\times n}$ est une matrice symétrique.
    - $*$ $\mathbf{x} \in \mathbb{R}^n$ est un vecteur réel.
    - $*$ $\mathbf{b} \in \mathbb{R}^n$ est un vecteur réel.
    - $*$ $c \in \mathbb{R}$ un scalaire réel.
  - Trouvez $\nabla^2_\mathbf{x} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla^2_\mathbf{x} f(\mathbf{x}) = \mathbf{A}$
    (b) $\nabla^2_\mathbf{x} f(\mathbf{x}) = \frac{1}{2}\mathbf{A}$
    (c) $\nabla^2_\mathbf{x} f(\mathbf{x}) = \mathbf{A} + \mathbf{A}^\mathrm{T}$
    (d) $\nabla^2_\mathbf{x} f(\mathbf{x}) = \mathbf{A}^\mathrm{T}\mathbf{A}$
  - Justifiez votre réponse ***(2.5 points)***.

- **Problème 1 : Question 3**
  - Soit $f(\mathbf{x}) = e^{\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}},$ où :
    - $*$ $\mathbf{A} \in \mathbb{R}^{n\times n}$ est une matrice symétrique.
    - $*$ $\mathbf{b} \in \mathbb{R}^n$ est un vecteur réel.
    - $*$ $c \in \mathbb{R}$ une constante réelle.
  - Trouvez $\nabla_\mathbf{x} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla f(\mathbf{x}) = e^{\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}}\mathbf{A}\mathbf{x}$
    (b) $\nabla f(\mathbf{x}) = 2e^{\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}}\mathbf{A}\mathbf{x}$
    (c) $\nabla f(\mathbf{x}) = e^{\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}}\mathbf{A}^\mathrm{T}\mathbf{x}$
    (d) $\nabla f(\mathbf{x}) = e^{\mathbf{x}^\mathrm{T}\mathbf{A}\mathbf{x}}(\mathbf{A}\mathbf{x} + \mathbf{A}^\mathrm{T}\mathbf{x})$
  - Justifiez votre réponse (Vous pouvez utiliser la règle de la dérivée en chaîne) ***(2.5 points)***.

- **Problème 1 : Question 4**
  - Soit $f(\mathbf{x}) = e^{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}$, où :
    - $\mathbf{A} \in \mathbb{R}^{n\times n}$ est une matrice symétrique.
    - $\mathbf{b} \in \mathbb{R}^{n}$ est un vecteur réel.
    - $c \in \mathbb{R}$ une constante réelle.
  - Trouvez $\nabla^2_{\mathbf{x}} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = e^{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}(2\mathbf{A} + 4\mathbf{A}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{A})$
    (b) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = e^{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}(\mathbf{A} + 4\mathbf{A}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{A})$
    (c) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = e^{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}(2\mathbf{A} + 2\mathbf{A}^{\mathrm{T}})$
    (d) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = e^{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}(2\mathbf{A} + 2\mathbf{A}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathbf{A})$
  - Justifiez votre réponse (Vous pouvez utiliser la règle de la dérivée du produit matriciel) ***(2.5 points)***.
- **Problème 1 : Question 5**
  - Soit $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\|^2 = (\mathbf{A}\mathbf{x})^{\mathrm{T}}(\mathbf{A}\mathbf{x})$, où :
    - $A \in \mathbb{R}^{m\times n}$ est une matrice donnée.
    - $\mathbf{x} \in \mathbb{R}^{n}$ est un vecteur réel composé de n éléments.
  - Trouvez $\nabla_{\mathbf{x}} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla_{\mathbf{x}} f(\mathbf{x}) = 2\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}$
    (b) $\nabla_{\mathbf{x}} f(\mathbf{x}) = \mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}$
    (c) $\nabla_{\mathbf{x}} f(\mathbf{x}) = 2\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{x}$
    (d) $\nabla_{\mathbf{x}} f(\mathbf{x}) = \mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{x}$
  - Justifiez votre réponse ***(2.5 points)***.
- **Problème 1 : Question 6**
  - Soit $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\|^2 = (\mathbf{A}\mathbf{x})^{\mathrm{T}}(\mathbf{A}\mathbf{x})$, où :
    - $A \in \mathbb{R}^{m\times n}$ est une matrice donnée.
    - $\mathbf{x} \in \mathbb{R}^{n}$ est un vecteur réel composé de n éléments.
  - Trouvez $\nabla^2_{\mathbf{x}} f(\mathbf{x})$ en choisissant la bonne réponse parmi les choix suivants ***(2.5 points)***:
    (a) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = 2\mathbf{A}^{\mathrm{T}}\mathbf{A}$
    (b) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = \mathbf{A}^{\mathrm{T}}\mathbf{A}$
    (c) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = 2\mathbf{A}\mathbf{A}^{\mathrm{T}}$
    (d) $\nabla^2_{\mathbf{x}} f(\mathbf{x}) = \mathbf{A}\mathbf{A}^{\mathrm{T}}$
  - Justifiez votre réponse ***(2.5 points)***.

2. **Problème 2 : Estimation par la Méthode du Maximum de Vraisemblance et Propriétés des Estimateurs** ***(30 points)***

- Soit $X_1, X_2, \dots, X_n$ un échantillon aléatoire issu d'une **loi de Poisson** de paramètre λ, c'est-à-dire :

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots$$

  (a) Déterminer l'estimateur du maximum de vraisemblance (EMV) du paramètre λ ***(5 points)***.
  (b) Vérifier si cet estimateur est biaisé ***(5 points)***.

(c) Vérifier si cet estimateur est consistant (ou convergent) en calculant sa variance ***(5 points)***.
On appelle **erreur quadratique moyenne** la quantité :

$$EQM(\hat{\theta}_n) = \mathbb{E}\left[(\hat{\theta}_n - \theta)^2\right] = V(\hat{\theta}_n) + [b(\hat{\theta}_n)]^2$$

Un estimateur est dit **consistant** si :

$$\lim_{n\to\infty} EQM(\hat{\theta}_n) = 0$$

.

(d) On observe les fréquences suivantes pour une variable X suivant une loi de Poisson :

x	$f_X(x)$
0	7
1	14
2	12
3	13
4	6
5	3

i. Trouver l'estimateur du maximum de vraisemblance (EMV) de λ en utilisant les données observées ***(5 points)***.

ii. Estimer la probabilité $P(X = 2)$ ***(5 points)***.

3. **Problème 3 : Test F de Fisher et Intervalle de Confiances** ***(20 points)***

- Une université teste une nouvelle méthode d'étude et souhaite comparer la variabilité des scores aux examens. Deux groupes d'étudiants ont passé le même examen :
  - **Groupe A (méthode classique) :** $[72, 75, 78, 69, 82, 77, 71, 74, 80, 73]$.
  - **Groupe B (nouvelle méthode) :** $[78, 81, 85, 79, 88, 84, 82, 80, 86, 83]$.
- Taille des échantillons : $n_A = n_B = 10$.

(a) Effectuer un **test F de Fisher** pour comparer les variances des deux groups pour un niveau de signification $\alpha = 0.05$ ***(5 points)***.

(b) Tracer la courbe de la distribution F montrant la région critique (zone de rejet de H_0) et la statistique F_{stat} observée ***(5 points)***.

(c) Calculer un intervalle de confiance pour le ratio des variances à 95% ***(5 points)***.

(d) Fournir le code source pour le calcul de l'intervalle de confiance et interpréter le résultat obtenu (Est-ce que la valeur d'un ratio égalr à 1 est dans l'intervalle?) ***(5 points)***.

4. **Problème 4 : Loi des Grands Nombres et Théorème Central Limite** ***(20 points)***

- On simule une variable aléatoire de Bernoulli ($p = 0.5$) et on calcule la moyenne empirique sur n observations. On vous fournit le code et on vous demande de modifier certains paramètres avant d'exécuter to observer les résultats.
- **Problème 4 - Question 1 :** Dans le code suivant, modifiez la valeur de $n = [50, 500, 5000]$. Quelle est l'influence de n sur la dispersion de $\overline{X}_n$ ***(5 points)***?

.

```python
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_theme()

np.random.seed(42)

# Paramètres
n = 100  # Nombre d'échantillons
num_trials = 2000  # Nombre de répétitions

# Génération des variables de Bernoulli
X = np.random.binomial(1, 0.5, (num_trials, n_max))

# Calcul des moyennes cumulatives
means = np.cumsum(X, axis=1) / np.arange(1, n_max + 1)

final_means = means[:, -1]  # Moyenne pour n = n_max

plt.figure(figsize=(10, 5))
plt.hist(final_means, bins=20, density=True, alpha=0.6, color='b', edgecolor='black')
plt.axvline(0.5, color='red', linestyle='dashed', label=r'$\mathbb{E}[X] = 0.5$')
plt.title("Loi faible des grands nombres : distribution de $\overline{X}_n$")
plt.xlabel("Valeur de la moyenne empirique")
plt.ylabel("Densité")
plt.legend()
plt.show()
```

- **Problème 4 - Question 2 :** Pour une valeur fixe de $n = 1000$, Que se passe-t-il si l'on répète l'expérience plusieurs fois ou même une seule fois ($num_{\text{trials}} = [1, 5, 10, 1000]$)? Qu'est-ce qui est observé en commun dans tous les cas de figures ***(10 points)***?

.

```python
# LOI FORTE DES GRANDS NOMBRES: Trajectoires individuelles
# Paramètres
n = 10  # Nombre d'échantillons
num_trials = 5  # Nombre de répétitions

# Génération des variables de Bernoulli
X = np.random.binomial(1, 0.5, (num_trials, n_max))

# Calcul des moyennes cumulatives
means = np.cumsum(X, axis=1) / np.arange(1, n_max + 1)

plt.figure(figsize=(10, 5))
for i in range(5):  # Tracer 5 trajectoires différentes
    plt.plot(means[i, :], alpha=0.7, label=f'Trajectoire {i+1}')

plt.axhline(0.5, color='red', linestyle='dashed', label=r'$\mathbb{E}[X] = 0.5$')
plt.title("Loi forte des grands nombres : trajectoires de $\overline{X}_n$")
plt.xlabel("Nombre d'échantillons $n$")
plt.ylabel("Moyenne empirique $\overline{X}_n$")
plt.legend()
plt.show()
```

- **Problème 4 - Question 3 :** Changez les valeurs de n $(100, 1000, 10000)$ et observez comment la distribution évolue. Vers quelle densité l'histogramme converge ***(5 points)***?

.

```python
import scipy.stats as stats

num_trials = 500
for n in [100, 1000, 10000]:
```

```python
X_tcl = np.random.binomial(1, 0.5, (num_trials, n))
sample_means = np.mean(X_tcl, axis=1)

plt.figure(figsize=(10, 5))
plt.hist(sample_means, bins=20, density=True, alpha=0.6, color='g', edgecolor='black', label="Moyennes empiriques")

# Superposition d'une distribution normale
mu, sigma = 0.5, np.sqrt(0.5 * 0.5 / n)
x_vals = np.linspace(0.3, 0.7, 100)
plt.plot(x_vals, stats.norm.pdf(x_vals, mu, sigma), color="red", lw=2, label="Densité normale")

plt.axvline(0.5, color='black', linestyle='dashed', label=r'$\mathbb{E}[X] = 0.5$')
plt.title(f"Théorème Central Limite : Distribution pour $n={n}$")
plt.xlabel("Valeur de la moyenne empirique")
plt.ylabel("Densité")
plt.legend()
plt.show()
```