**Date remise: Le 5 décembre, 23h**

Instructions

- *Montrez votre démarche pour toutes les questions !*
- *Utilisez LaTeX et le modèle que nous vous fournissons pour rédiger vos réponses. Vous pouvez réutiliser la plupart des raccourcis de notation, des équations et/ou des tableaux. SVP voir la politique des devoirs sur le site web du cours pour plus de détails.*
- *Vous devez soumettre toutes vos réponses sur la page Gradescope du cours*
- *Les TAs pour ce devoir sont* ***Arian Khorasani et Sarthak Mittal***

**Question 1** (5)**.** (**Divergence de Kullback-Leibler**)

Étant donné deux distributions gaussiennes univariées, $p(x) = \mathcal{N}(\mu_1, \sigma_1^2)$ et $q(x) = \mathcal{N}(\mu_2, \sigma_2^2)$, trouver la $\mathbb{KL}$-Divergence entre la distribution $q$ et la distribution $p$. En particulier, dériver l'expression de forme fermée pour

$$\mathbb{KL}[q(x)||p(x)] = \mathbb{E}_{q(x)}\left[\log \frac{q(x)}{p(x)}\right]$$

C'est le même que $\mathbb{KL}[p(x)||q(x)]$ ?

**Réponse :**
Calcul de la $\mathbb{KL}$-Divergence entre la distribution $q$ et la distribution $p$:

$$
\begin{aligned}
\mathbb{KL}[q(x)||p(x)] &= \mathbb{E}_{q(x)}\left[\log \frac{q(x)}{p(x)}\right] \\
&= \mathbb{E}_{q(x)}[\log q(x)] - \mathbb{E}_{q(x)}[\log p(x)] \\
&= \mathbb{E}_{q(x)}[-\frac{1}{2}log(2\pi\sigma_2^2) - \frac{(x-\mu_2)^2}{2\sigma_2^2}] - \mathbb{E}_{q(x)}[-\frac{1}{2}log(2\pi\sigma_1^2) - \frac{(x-\mu_1)^2}{2\sigma_1^2}] \\
&= -\mathbb{E}_{q(x)}[\frac{1}{2}log(2\pi\sigma_2^2)] - \mathbb{E}_{q(x)}[\frac{(x-\mu_2)^2}{2\sigma_2^2}] + \mathbb{E}_{q(x)}[\frac{1}{2}log(2\pi\sigma_1^2)] + \mathbb{E}_{q(x)}[\frac{(x-\mu_1)^2}{2\sigma_1^2}] \\
&= \mathbb{E}_{q(x)}[\frac{1}{2}log(2\pi\sigma_1^2)] - \mathbb{E}_{q(x)}[\frac{1}{2}log(2\pi\sigma_2^2)] + \mathbb{E}_{q(x)}[\frac{(x-\mu_1)^2}{2\sigma_1^2}] - \mathbb{E}_{q(x)}[\frac{(x-\mu_2)^2}{2\sigma_2^2}] \\
&= \frac{1}{2}log(2\pi\sigma_1^2) - \frac{1}{2}log(2\pi\sigma_2^2) + \mathbb{E}_{q(x)}[\frac{(x-\mu_1)^2}{2\sigma_1^2}] - \mathbb{E}_{q(x)}[\frac{(x-\mu_2)^2}{2\sigma_2^2}] \\
&= \frac{1}{2}log(\frac{2\pi\sigma_1^2}{2\pi\sigma_2^2}) + \mathbb{E}_{q(x)}[\frac{(x-\mu_1)^2}{2\sigma_1^2}] - \mathbb{E}_{q(x)}[\frac{(x-\mu_2)^2}{2\sigma_2^2}] \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \mathbb{E}_{q(x)}[\frac{(x-\mu_1)^2}{2\sigma_1^2}] - \mathbb{E}_{q(x)}[\frac{(x-\mu_2)^2}{2\sigma_2^2}] \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}\mathbb{E}_{q(x)}[(x-\mu_1)^2] - \frac{1}{2\sigma_2^2}\mathbb{E}_{q(x)}[(x-\mu_2)^2] \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}\mathbb{E}_{q(x)}[x^2 - 2x\mu_1 + \mu_1^2] - \frac{1}{2\sigma_2^2}\mathbb{E}_{q(x)}[x^2 - 2x\mu_2 + \mu_2^2] \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}(\mathbb{E}_{q(x)}[x^2] - 2\mu_1\mathbb{E}_{q(x)}[x] + \mathbb{E}_{q(x)}[\mu_1^2]) - \frac{1}{2\sigma_2^2}(\mathbb{E}_{q(x)}[x^2] - 2\mu_2\mathbb{E}_{q(x)}[x] + \mathbb{E}_{q(x)}[\mu_2^2])
\end{aligned}
$$

Or :

$$
\begin{aligned}
\mathbb{E}_{q(x)}[x^2] &= \sigma_2^2 + \mu_2^2 \\
\mathbb{E}_{q(x)}[x] &= \mu_2 \\
\mathbb{E}_{q(x)}[\mu_1^2] &= \mu_1^2 \\
\mathbb{E}_{q(x)}[\mu_2^2] &= \mu_2^2
\end{aligned}
$$

D'où :

$$
\begin{aligned}
\mathbb{KL}[q(x)||p(x)] &= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}(\mathbb{E}_{q(x)}[x^2] - 2\mu_1\mathbb{E}_{q(x)}[x] + \mathbb{E}_{q(x)}[\mu_1^2]) - \frac{1}{2\sigma_2^2}(\mathbb{E}_{q(x)}[x^2] - 2\mu_2\mathbb{E}_{q(x)}[x] + \mathbb{E}_{q(x)}[\mu_2^2]) \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}(\sigma_2^2 + \mu_2^2 - 2\mu_1\mu_2 + \mu_1^2) - \frac{1}{2\sigma_2^2}(\sigma_2^2 + \mu_2^2 - 2\mu_2^2 + \mu_2^2) \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}(\sigma_2^2 + (\mu_2 - \mu_1)^2) - \frac{\sigma_2^2}{2\sigma_2^2} \\
&= \frac{1}{2}log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{1}{2\sigma_1^2}(\sigma_2^2 + (\mu_2 - \mu_1)^2) - \frac{1}{2} \\
&= \frac{1}{2}(log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_1^2} - 1)
\end{aligned}
$$

Donc :

$$\boxed{\mathbb{KL}[q(x)||p(x)] = \frac{1}{2}(log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_1^2} - 1)}$$

Généralement $\mathbb{KL}[p(x)||q(x)]$ n'est pas le même que $\mathbb{KL}[q(x)||p(x)]$.

Si $\sigma_1 \neq \sigma_2$ :

$$\mathbb{KL}[p(x)||q(x)] \neq \mathbb{KL}[q(x)||p(x)]$$

En effet, dans ce cas précis :

$$\frac{1}{2}(log(\frac{\sigma_2^2}{\sigma_1^2}) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{\sigma_2^2} - 1) \neq \frac{1}{2}(log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_1^2} - 1)$$

Puisque :

$$\frac{\sigma_2^2}{\sigma_1^2} \neq \frac{\sigma_1^2}{\sigma_2^2}$$

et

$$\frac{(\mu_1 - \mu_2)^2}{\sigma_1^2} \neq \frac{(\mu_1 - \mu_2)^2}{\sigma_2^2}$$

Or, si $\sigma_1 = \sigma_2$, on a bien :

$$\begin{aligned}
\mathbb{KL}[q(x)||p(x)] &= \frac{1}{2}(log(\frac{\sigma_1^2}{\sigma_2^2}) + \frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_1^2} - 1) \\
&= \frac{1}{2}(\frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_1^2} - 1) \\
&= \frac{1}{2}(\frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{1}{2}(1 + \frac{(\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{(\mu_2 - \mu_1)^2}{2\sigma_2^2}
\end{aligned}$$

et, puisque $\sigma_1 = \sigma_2$:

$$\begin{aligned}
\mathbb{KL}[p(x)||q(x)] &= \frac{1}{2}(log(\frac{\sigma_2^2}{\sigma_1^2}) + \frac{\sigma_1^2 + (\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{1}{2}(\frac{\sigma_1^2 + (\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{1}{2}(\frac{\sigma_2^2 + (\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{1}{2}(1 + \frac{(\mu_2 - \mu_1)^2}{\sigma_2^2} - 1) \\
&= \frac{(\mu_2 - \mu_1)^2}{2\sigma_2^2}
\end{aligned}$$

On a alors, si $\sigma_1 = \sigma_2$ :

$$\mathbb{KL}[p(x)||q(x)] = \mathbb{KL}[q(x)||p(x)]$$

**Question 2** (5-5-5-6)**.** (**Les flux normalisants**)

Les flux normalisants (*normalizing flows*) sont des transformations inversibles et expressives des lois de probabilités.

Dans cet exercice, nous allons explorer la possibilité d'inverser quelques transformations. Dans les 3 premières questions, nous considérons une fonction déterministe $g : z \in \mathbb{R} \to \mathbb{R}$.

1. Soient $g(z) = af(bz + c)$ et $f$ est un redresseur ( *rectified linear unit* ) $f(x) = \max(0, x)$. Montrez que la fonction $g$ n'est pas inversible.
2. Soit $g : z \in \mathbb{R} \mapsto \mathbb{R}$, où $g(z) = \sigma^{-1}(\sum_{i=1}^{N} w_i \sigma(a_i z + b_i))$, $0 < w_i < 1$, $\sum_i w_i = 1$, et $a_i > 0$. Montrez que $g$ est *strictement croissante* sur son domaine de définition $(-\infty, \infty)$ [1]
3. Soit $g(z) = z + f(z)$ et $df/dz > -1$. Montrez que $g$ est inversible.
4. Considérez la transformation suivante

$$g(\boldsymbol{z}) = \boldsymbol{z} + \beta h(\alpha, r)(\boldsymbol{z} - \boldsymbol{z}_0) \tag{1}$$

où $\boldsymbol{z}_0 \in \mathbb{R}^D$, $\alpha \in \mathbb{R}^+$, $\beta \in \mathbb{R}$, et $r = ||\boldsymbol{z} - \boldsymbol{z}_0||_2$, $h(\alpha, r) = 1/(\alpha + r)$. Considérez également la décomposition suivante $\boldsymbol{z} = \boldsymbol{z}_0 + r\tilde{\boldsymbol{z}}$. (i) Étant donné $\boldsymbol{y} = g(\boldsymbol{z})$, montrez que $\beta \geq -\alpha$ est une condition suffisante pour obtenir une valeur unique de $r$ à partir de l'équation (1). (ii) Étant donnés $r$ et $\boldsymbol{y}$, montrez que l'équation (1) possède une unique solution $\tilde{\boldsymbol{z}}$.

1. Pour écrire votre réponse à cette question, rappelez vous que si une fonction $f$ est *strictement croissante*, alors elle est injective sur son domaine de définition. De plus, si $T$ est l'image de la fonction $f$, alors $f$ possède une fonction inverse $f^{-1}$ sur $T$. Vous pouvez considérer une fonction *strictement croissante*, *i.e.* $df(x)/dx > 0$, par exemple.

**Réponse :**

1. Soient $g(z) = af(bz + c)$ et $f$ est un redresseur ( *rectified linear unit* ) $f(x) = \max(0, x)$.
   Montrons que la fonction g n'est pas inversible.
   Une fonction est inversible si à tout élément de l'ensemble de départ correspond un unique élément de l'ensemble d'arrivée, et si tout élément de l'ensemble d'arrivée est l'image d'un unique élément de l'ensemble de départ.
   Autrement dit, une fonction $g$ est inversible si et seulement si $g$ est injective et surjective (donc bijective).
   Montrons ici que $g$ n'est pas injective, $\forall\ a, b, c$ et $z \in \mathbb{R}$ .
   En d'autres termes, montrons qu'il existe au moins deux éléments distincts de l'ensemble de départ peuvent avoir la même image par $g$.

   Pour se faire, nous allons faire plusieurs disjonctions de cas, et trouver à chaque fois des contres exemples tel que $z \neq z'$ et $g(z) = g(z')$ :

   (a) **Cas : $a = 0$**
   Si $a = 0$, alors $g(z) = 0$, $\forall z \in \mathbb{R}$. Donc $g$ n'est dans ce cas pas bijective et n'est pas inversible.

   (b) **Cas : $a \neq 0$**
   Faisons une deuxième disjonction de cas selon $b$.

   i. **Cas : $b = 0$**
   Si $b = 0$, on a :
   $$g(z) = af(bz + c) = af(c)$$
   Ainsi, on a $\forall\ z \in \mathbb{R}$, $g(z) = a\max(0, c)$. Donc $g$ n'est dans ce cas pas bijective et n'est pas inversible.

   ii. **Cas : $b \neq 0$**
   Divisons une dernière fois en plusieurs cas en fonction de $c$ :

   A. **Cas : $c > 0$**
   Considérons $z = \frac{-c}{b}$ et $z' = \frac{-2c}{b}$ deux éléments de l'ensemble de départ de $g$.
   On a alors, avec $c > 0$:
   $$\begin{aligned} g(z) &= af(b\frac{-c}{b} + c) \\ &= af(0) \\ &= a\max(0, 0) \\ &= 0 \end{aligned}$$
   et
   $$\begin{aligned} g(z') &= af(b\frac{-2c}{b} + c) \\ &= af(-c) \\ &= a\max(0, -c) \\ &= 0 \end{aligned}$$

On a donc $g(z) = g(z')$ avec $z \neq z'$, donc $g$ n'est dans ce cas pas bijective et n'est donc pas inversible.

B. **Cas : $c < 0$**
Considérons $z = \frac{c}{b}$ et $z' = \frac{2c}{b}$ deux éléments de l'ensemble de départ de $g$.
On a alors, avec $c < 0$:

$$\begin{aligned} g(z) &= af(b\frac{c}{b} + c) \\ &= af(2c) \\ &= a\max(0, 2c) \\ &= 0 \end{aligned}$$

et

$$\begin{aligned} g(z') &= af(b\frac{2c}{b} + c) \\ &= af(3c) \\ &= a\max(0, 3c) \\ &= 0 \end{aligned}$$

On a donc encore une fois $g(z) = g(z')$ avec $z \neq z'$, donc $g$ n'est dans ce cas pas bijective et n'est donc pas inversible.

C. **Cas : $c = 0$**
Enfin si $c = 0$ on a :

$$g(z) = af(bz) = a\max(0, bz)$$

Cette fois-ci, si on considère $z = \frac{-1}{b}$ et $z' = \frac{-2}{b}$ deux éléments de l'ensemble de départ de $g$, on a alors :

$$\begin{aligned} g(z) &= af(b\frac{-1}{b}) \\ &= af(-1) \\ &= a\max(0, -1) \\ &= 0 \end{aligned}$$

et

$$\begin{aligned} g(z') &= af(b\frac{-2}{b}) \\ &= af(-2) \\ &= a\max(0, -2) \\ &= 0 \end{aligned}$$

On a donc encore une nouvelle fois $g(z) = g(z')$ avec $z \neq z'$, donc $g$ n'est dans ce cas pas bijective et n'est donc pas inversible.

Ainsi, dans tous les cas, $\forall\ a, b, c$ et $z \in \mathbb{R}$, on a que $g$ n'est pas inversible.

2. Soit $g : z \in \mathbb{R} \mapsto \mathbb{R}$, où $g(z) = \sigma^{-1}(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i))$, $0 < w_i < 1$, $\sum_i w_i = 1$, et $a_i > 0$. Montrons que $g$ est *strictement croissante* sur son domaine de définition $(-\infty, \infty)$:
Sachant que $\sigma^{-1}(X) = log(\frac{X}{1-X})$ $,\forall\ X$ :

$$
\begin{aligned}
g(z) &= \sigma^{-1}(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i)) \\
&= log(\frac{\sum_{i=1}^{N} w_i\sigma(a_i z + b_i)}{1 - \sum_{i=1}^{N} w_i\sigma(a_i z + b_i)}) \\
&= log(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i)) - log(1 - \sum_{i=1}^{N} w_i\sigma(a_i z + b_i))
\end{aligned}
$$

Pour montrer que $g$ est *strictement croissante* sur son domaine de définition, regardons le signe de sa dérivée $\frac{d}{dz}g$, notée $g'$ :

$$
g'(z) = \frac{d}{dz}log(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i)) - \frac{d}{dz}log(1 - \sum_{i=1}^{N} w_i\sigma(a_i z + b_i))
$$

Calculons $\frac{d}{dz}log(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i))$ et $\frac{d}{dz}log(1 - \sum_{i=1}^{N} w_i\sigma(a_i z + b_i))$ :

$$
\frac{d}{dz}log(\sum_{i=1}^{N} w_i\sigma(a_i z + b_i)) = \frac{\sum_{i=1}^{N} a_i w_i \frac{e^{a_i z + b_i}}{(1+e^{a_i z + b_i})^2}}{\sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z + b_i}}}
$$

De plus :

$$
\frac{d}{dz}log(1 - \sum_{i=1}^{N} w_i\sigma(a_i z + b_i)) = \frac{-\sum_{i=1}^{N} a_i w_i \frac{e^{a_i z + b_i}}{(1+e^{a_i z + b_i})^2}}{1 - \sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z + b_i}}}
$$

Ainsi :

$$
g'(z) = \frac{\sum_{i=1}^{N} a_i w_i \frac{e^{a_i z + b_i}}{(1+e^{a_i z + b_i})^2}}{\sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z + b_i}}} + \frac{\sum_{i=1}^{N} a_i w_i \frac{e^{a_i z + b_i}}{(1+e^{a_i z + b_i})^2}}{1 - \sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z + b_i}}}
$$

Or, avec $a_i > 0$, $0 < w_i < 1$, $\forall\ i$ on a :

$$
\sum_{i=1}^{N} a_i w_i \frac{e^{a_i z + b_i}}{(1 + e^{a_i z + b_i})^2} > 0
$$

et

$$
\sum_{i=1}^{N} w_i \frac{1}{1 + e^{a_i z + b_i}} > 0
$$

Qu'en est-il du signe de $1 - \sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z+b_i}}$ ?
On sait que, $\forall\ i$ on a :

$$\begin{aligned} e^{a_i z+b_i} &> 0 \\ 1+e^{a_i z+b_i} &> 1 \\ \frac{1}{1+e^{a_i z+b_i}} &< 1 \end{aligned}$$

En utilisant le fait que $0 < w_i < 1$, $\forall\ i$ on a alors :

$$w_i \frac{1}{1+e^{a_i z+b_i}} < 1$$

De plus, sachant que $\sum_{i=1}^{N} w_i = 1$ et que $0 < \frac{1}{1+e^{a_i z+b_i}} < 1$, alors :

$$\sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z+b_i}} < 1$$

D'où :

$$1 - \sum_{i=1}^{N} w_i \frac{1}{1+e^{a_i z+b_i}} > 0$$

Donc on a bien :

$$\begin{aligned} \frac{d}{dz} log(\sum_{i=1}^{N} w_i \sigma(a_i z + b_i)) &> 0 \\ -\frac{d}{dz} log(1 - \sum_{i=1}^{N} w_i \sigma(a_i z + b_i)) &> 0 \end{aligned}$$

et

$$g'(z) = \frac{d}{dz} log(\sum_{i=1}^{N} w_i \sigma(a_i z + b_i)) - \frac{d}{dz} log(1 - \sum_{i=1}^{N} w_i \sigma(a_i z + b_i)) > 0$$

Avec $g'(z)$ strictement positive, on a bien que $g$ est strictement croissante sur son domaine de définition.

3. Le théorème des valeurs intermédiaires et son corollaire, le théorème de la bijection, assurent que toute application **continue strictement monotone** sur un intervalle I détermine une bijection de I sur f(I) = J et que J est aussi un intervalle.
   Cela signifie qu'une telle fonction est inversible et possède une application réciproque définie sur J à valeurs dans I.
   On considère ici $g(z) = z + f(z)$ avec $df/dz > -1$.
   On a alors : $g'(z) = 1 + f'(z)$.
   Or $f'(z) > -1$, donc $g'(z) > 0$.
   Ainsi, $g'(z) > 0$, et $g$ est strictement croissante (donc strictement monotone). Donc $g$ est inversible.

4. (a) On a :

$$y = \boldsymbol{z}_0 + r\tilde{\boldsymbol{z}} + \frac{\beta r\tilde{\boldsymbol{z}}}{\alpha + r}$$
$$y = \boldsymbol{z}_0 + r\tilde{\boldsymbol{z}} + \frac{\beta\tilde{\boldsymbol{z}}}{\frac{\alpha}{r} + 1}$$
$$y(\frac{\alpha}{r} + 1) = \boldsymbol{z}_0(\frac{\alpha}{r} + 1) + r\tilde{\boldsymbol{z}}(\frac{\alpha}{r} + 1) + \beta\tilde{\boldsymbol{z}}$$
$$(y - \boldsymbol{z}_0)\frac{\alpha}{r} = \beta\tilde{\boldsymbol{z}} + \tilde{\boldsymbol{z}}\alpha + r\tilde{\boldsymbol{z}} + \boldsymbol{z}_0 - y$$
$$\alpha(y - \boldsymbol{z}_0) = r\beta\tilde{\boldsymbol{z}} + r\tilde{\boldsymbol{z}}\alpha + r^2\tilde{\boldsymbol{z}} + r\boldsymbol{z}_0 - ry$$
$$\tilde{\boldsymbol{z}}r^2 + (\beta\tilde{\boldsymbol{z}} + \tilde{\boldsymbol{z}}\alpha + \boldsymbol{z}_0 - y)r - \alpha(y - \boldsymbol{z}_0) = 0$$

On a alors un équation avec un polynôme du seconde degré à résoudre.
Pour que l'équation ait une unique solution r, il faut que le déterminant $\Delta = 0$.
Il faut donc montrer que :

$$\Delta = 0$$
$$(\beta\tilde{\boldsymbol{z}} + \tilde{\boldsymbol{z}}\alpha + \boldsymbol{z}_0 - y)^2 + 4\tilde{\boldsymbol{z}}\alpha(y - \boldsymbol{z}_0) = 0$$

(b) Pour $r$ et $\boldsymbol{y}$ donnés, montrons que l'équation (1) possède une unique solution $\tilde{\boldsymbol{z}}$ :
On a :

$$g(\boldsymbol{z}) = \boldsymbol{z} + \beta h(\alpha, r)(\boldsymbol{z} - \boldsymbol{z}_0)$$
$$y = \boldsymbol{z}_0 + r\tilde{\boldsymbol{z}} + \beta \times \frac{r\tilde{\boldsymbol{z}}}{\alpha + r}$$
$$y - \boldsymbol{z}_0 = \tilde{\boldsymbol{z}}(r + \frac{\beta r}{\alpha + r})$$
$$\tilde{\boldsymbol{z}} = \frac{y - \boldsymbol{z}_0}{r + \frac{\beta r}{\alpha + r}}$$

Ainsi, pour $r$ et $\boldsymbol{y}$ donnés, l'équation (1) possède une unique solution $\boxed{\tilde{\boldsymbol{z}} = \frac{y - \boldsymbol{z}_0}{r + \frac{\beta r}{\alpha + r}}}$.

**Question 3** (3-5-2-2-3)**.** (**Combiné Auto-encodeur variationnel et modèle de diffusion**) DDPMs sont une classe de modèles génératifs qui s'appuient sur un processus de diffusion directe connu. $q(x_t|x_{t-1})$, qui détruit progressivement la structure des données jusqu'à ce qu'elle converge au bruit non structuré, par exemple. $\mathcal{N}(0, I)$ et un processus paramétré appris (par un réseau neuronal !) en arrière $p_\theta(x_{t-1}|x_t)$ qui élimine le bruit de façon itérative jusqu'à ce que vous ayez obtenu un échantillon de la distribution de données.

Soit $\boldsymbol{x}_0$ être un échantillon de la distribution des données ; et laisser le processus de diffusion directe (processus bruyant) être défini en utilisant

$$q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0) := \prod_{t=1}^{T} q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) \qquad \text{where} \qquad q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \mathcal{N}(\boldsymbol{x}_t; \sqrt{1-\beta_t}\boldsymbol{x}_{t-1}, \beta_t I)$$

et le processus de diffusion inverse (processus de dénotation) étant un processus appris suivant

$$p_\phi(\boldsymbol{x}_{0:T}) = p(\boldsymbol{x}_T)\prod_{t=1}^{T} p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) \qquad \text{where} \qquad p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) = \mathcal{N}(\boldsymbol{x}_{t-1}|\boldsymbol{\mu}_\phi(\boldsymbol{x}_t, t), \Sigma_\phi(\boldsymbol{x}_t, t))$$

(a) Montre ça $\log p_\phi(\boldsymbol{x}_0) \geq \underbrace{\mathbb{E}_q\left[\log \frac{p_\phi(\boldsymbol{x}_{0:T})}{q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)}\right]}_{\mathcal{L}_{DDPM}}$ et puis, montre ça

$$\mathbb{E}_q\left[\log \frac{p_\phi(\boldsymbol{x}_{0:T})}{q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)}\right] = \mathbb{E}\left[\log p(\boldsymbol{x}_T) + \sum_{t=1}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right]$$

(b) Montre ça $\mathcal{L}_{DDPM} = \mathbb{E}_q\left[\log p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1) - \sum_{t=2}^{T} \mathbb{KL}\left[q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0) \mid\mid p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)\right] - \mathbb{KL}[q(\boldsymbol{x}_T|\boldsymbol{x}_0) \mid\mid p(\boldsymbol{x}_T)]\right]$

(c) Considérons maintenant un échantillon de données du formulaire $(\boldsymbol{x}_0, \boldsymbol{c})$, où $\boldsymbol{c}$ est maintenant des données auxiliaires supplémentaires qui vous ont été fournies (en particulier ; $\boldsymbol{c}$ peut être le même que $\boldsymbol{x}_0$ aussi bien). Supposons que nous modélisons les données comme suit $p(\boldsymbol{x}_{0:T}, \boldsymbol{c}, \boldsymbol{z})$ où nos variables latentes sont maintenant $\boldsymbol{z}$ et $\boldsymbol{x}_{1:T}$. Puisque le postérieur sera intraitable, essayons de l'approcher avec $q(\boldsymbol{x}_{1:T}, \boldsymbol{z}|\boldsymbol{c}, \boldsymbol{x}_0)$. Pouvez-vous maintenant re-dériver l'ELBO, qui est la limite inférieure de la probabilité logarithmique, comme $\log p(\boldsymbol{x}_0, \boldsymbol{c}) \geq \mathbb{E}_{q(\boldsymbol{x}_{1:T}, \boldsymbol{z})}\left[\log \frac{p(\boldsymbol{x}_{0:T}, \boldsymbol{c}, \boldsymbol{z})}{q(\boldsymbol{x}_{1:T}, \boldsymbol{z}|\boldsymbol{x}_0, \boldsymbol{c})}\right]$ ?

(d) Supposons maintenant que vous modélisez ce problème avec une combinaison de VAE et Modèle probabiliste de diffusion de débruitage, où l'encodeur du VAE a les paramètres $\psi$, le décodeur $\theta$ et le modèle de débruitage $\phi$. Dans ce cas, la distribution générative se factorise comme

$$\begin{aligned} p(\boldsymbol{x}_{0:T}, \boldsymbol{c}, \boldsymbol{z}) &= p(\boldsymbol{z})p_\theta(\boldsymbol{c}|\boldsymbol{z})p_\phi(\boldsymbol{x}_{0:T}|\boldsymbol{c}, \boldsymbol{z}) \\ &= p(\boldsymbol{z})p_\theta(\boldsymbol{c}|\boldsymbol{z})p(\boldsymbol{x}_T|\boldsymbol{c}, \boldsymbol{z})\prod_{t=1}^{T} p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t\boldsymbol{c}, \boldsymbol{z}) \end{aligned}$$

En outre, supposons que nous voulons maintenant modéliser $\boldsymbol{z}$ utilisant un encodeur VAE avec paramètres $\psi$, et ensuite les autres variables latentes $\boldsymbol{x}_{1:T}$ subordonnée à $\boldsymbol{z}$ par le processus de diffusion directe. Pouvez-vous fournir une factorisation de $q(\boldsymbol{x}_{1:T}, \boldsymbol{z}|\boldsymbol{c}, \boldsymbol{x}_0)$ qui respecte cela ?

(e) Par la manipulation arithmétique de l'ELBO vous avez dérivé ci-dessus ainsi que la factorisation que vous avez fournies, Pouvez-vous maintenant décomposer l'objectif en une composante VAE et une composante DDPM ?

**Réponse :**

(a) Montrons que $\log p_\phi(\boldsymbol{x}_0) \geq \mathbb{E}_q\left[\log \frac{p_\phi(\boldsymbol{x}_{0:T})}{q(\boldsymbol{x}_{1:T}|\boldsymbol{x}_0)}\right]$.

$$\begin{aligned}
\log p_\phi(x_0) &= log \int_{x_{1:T}} p_\phi(x_0, x_{1:T})\, d(x_{1:T}) \\
&= \log \int_{x_{1:T}} p_\phi(x_0, x_{1:T}) \frac{q(x_{1:T}|x_0)}{q(x_{1:T}|x_0)}\, d(x_{1:T}) \\
&= \log \mathbb{E}_{x_{1:T} \sim q_\theta(x_{1:T}|x_0)} \left[\frac{p_\phi(x_0, x_{1:T})}{q(x_{1:T}|x_0)}\right] \\
&= \log \mathbb{E}_q \left[\frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right]
\end{aligned}$$

Or d'après l'inégalité de Jensen :

$$\begin{aligned}
\log p_\phi(x_0) &= \log \mathbb{E}_q \left[\frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right] \\
&\geq \mathbb{E}_q \left[\log \frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right]
\end{aligned}$$

Donc :

$$\boxed{\log p_\phi(x_0) \geq \mathbb{E}_q \left[\log \frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right]}$$

De plus, $\mathbb{E}_q\left[\log \frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right] = \mathbb{E}_q\left[\log \frac{p(\boldsymbol{x}_T)\prod_{t=1}^T p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{\prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right]$.
Or, d'après les propriétés du logarithme : $log(ab) = log(a) + log(b)$, on a :

$$\begin{aligned}
\mathbb{E}_q \left[\log \frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right] &= \mathbb{E}_q \left[\log \frac{p(\boldsymbol{x}_T)\prod_{t=1}^T p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{\prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right] \\
&= \mathbb{E}_q \left[\log p(\boldsymbol{x}_T) + \sum_{t=1}^T \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right]
\end{aligned}$$

Ainsi :

$$\boxed{\mathbb{E}_q \left[\log \frac{p_\phi(x_{0:T})}{q(x_{1:T}|x_0)}\right] = \mathbb{E}_q \left[\log p(\boldsymbol{x}_T) + \sum_{t=1}^T \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right]}$$

(b) Montrons que :

$$\mathcal{L}_{DDPM} = \mathbb{E}_q \left[\log p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1) - \sum_{t=2}^T \mathbb{KL}\left[q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0) \mid\mid p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)\right] - \mathbb{KL}[q(\boldsymbol{x}_T|\boldsymbol{x}_0) \mid\mid p(\boldsymbol{x}_T)]\right]$$

D'après la question précédente, on a :

$$\begin{aligned}
\mathcal{L}_{DDPM} &= \mathbb{E}_q \left[\log p(\boldsymbol{x}_T) + \sum_{t=1}^T \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}\right] \\
&= \mathbb{E}_q \left[\log p(\boldsymbol{x}_T) + \sum_{t=2}^T \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})} + \log \frac{p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1)}{q(\boldsymbol{x}_1|\boldsymbol{x}_0)}\right]
\end{aligned}$$

Or, en utilisant le théorème de Bayes et l'hypothèse de Markov :

$$\begin{aligned}\mathcal{L}_{DDPM} &= \mathbb{E}_q\left[\log p(\boldsymbol{x}_T) + \sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})} + \log \frac{p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1)}{q(\boldsymbol{x}_1|\boldsymbol{x}_0)}\right] \\ &= \mathbb{E}_q\left[\log p(\boldsymbol{x}_T) + \sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)} \cdot \frac{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0)}{q(\boldsymbol{x}_t|\boldsymbol{x}_0)} + \log \frac{p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1)}{q(\boldsymbol{x}_1|\boldsymbol{x}_0)}\right]\end{aligned}$$

Or :

$$\sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)} \cdot \frac{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0)}{q(\boldsymbol{x}_t|\boldsymbol{x}_0)} = \sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)} + \sum_{t=2}^{T} \log \frac{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0)}{q(\boldsymbol{x}_t|\boldsymbol{x}_0)}$$

De plus, on a que :

$$\begin{aligned}\sum_{t=2}^{T} \log \frac{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0)}{q(\boldsymbol{x}_t|\boldsymbol{x}_0)} &= \sum_{t=2}^{T} \log q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_0) - \sum_{t=2}^{T} \log q(x_t|\boldsymbol{x}_0) \\ &= \sum_{t=1}^{T-1} \log q(\boldsymbol{x}_t|\boldsymbol{x}_0) - \sum_{t=2}^{T} \log q(\boldsymbol{x}_t|\boldsymbol{x}_0) \\ &= \log q(\boldsymbol{x}_1|\boldsymbol{x}_0) + \sum_{t=2}^{T-1} \log q(\boldsymbol{x}_t|\boldsymbol{x}_0) - \sum_{t=2}^{T-1} \log q(\boldsymbol{x}_t|\boldsymbol{x}_0) - \log q(\boldsymbol{x}_T|\boldsymbol{x}_0) \\ &= \log q(\boldsymbol{x}_1|\boldsymbol{x}_0) - \log q(\boldsymbol{x}_T|\boldsymbol{x}_0)\end{aligned}$$

On a alors :

$$\begin{aligned}\mathcal{L}_{DDPM} &= \mathbb{E}_q\left[\log p(\boldsymbol{x}_T) + \sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)} + \log q(\boldsymbol{x}_1|x_0) - \log q(\boldsymbol{x}_T|\boldsymbol{x}_0) + \log \frac{p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1)}{q(\boldsymbol{x}_1|\boldsymbol{x}_0)}\right] \\ &= \mathbb{E}_q\left[\log \frac{p(\boldsymbol{x}_T)}{q(\boldsymbol{x}_T|\boldsymbol{x}_0)} + \sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)} + \log p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1)\right]\end{aligned}$$

Or :

$$\mathbb{E}_q\left[\log \frac{p(\boldsymbol{x}_T)}{q(\boldsymbol{x}_T|\boldsymbol{x}_0)}\right] = -\mathbb{KL}\left[q(\boldsymbol{x}_T|\boldsymbol{x}_0) \mid\mid p(\boldsymbol{x}_T)\right]$$

$$\mathbb{E}_q\left[\sum_{t=2}^{T} \log \frac{p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)}{q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0)}\right] = -\sum_{t=2}^{T} \mathbb{KL}\left[q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, x_0) \mid\mid p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)\right]$$

On a donc :

$$\boxed{\mathcal{L}_{DDPM} = \mathbb{E}_q\left[\log p_\phi(\boldsymbol{x}_0|\boldsymbol{x}_1) - \sum_{t=2}^{T} \mathbb{KL}\left[q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0) \mid\mid p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)\right] - \mathbb{KL}[q(\boldsymbol{x}_T|\boldsymbol{x}_0) \mid\mid p(\boldsymbol{x}_T)]\right]}$$

(c) On a ici :

$$\begin{aligned}
\mathcal{L}_{DDPM} &= \mathbb{E}_{q(\boldsymbol{x}_{1:T}\,,\,\boldsymbol{z})}\left[\log \frac{p(\boldsymbol{x}_{0:T}\,,\,\boldsymbol{c}\,,\,\boldsymbol{z})}{q(\boldsymbol{x}_{1:T}\,,\,\boldsymbol{z}|\boldsymbol{x}_0\,,\,\boldsymbol{c})}\right] \\
&= \mathbb{E}_q\left[\log \frac{p(\boldsymbol{z})p_\theta(\boldsymbol{c}|\boldsymbol{z})p(\boldsymbol{x}_T|\boldsymbol{c},\boldsymbol{z})\prod_{t=1}^T p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t\boldsymbol{c},\boldsymbol{z})}{q(\boldsymbol{z}|\boldsymbol{x}_{1:T})\ \prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{c},\boldsymbol{x}_0)}\right] \\
&= \mathbb{E}_q\left[\log \frac{p(\boldsymbol{z})p_\theta(\boldsymbol{c}|\boldsymbol{z})p(\boldsymbol{x}_T|\boldsymbol{c},\boldsymbol{z})\prod_{t=1}^T p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t\boldsymbol{c},\boldsymbol{z})}{q(\boldsymbol{z}|\boldsymbol{x}_{1:T})\ \prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{c},\boldsymbol{x}_0)}\right] \\
&= \mathbb{E}_q[\log p(\boldsymbol{z}) + \log p_\theta(\boldsymbol{c}|\boldsymbol{z}) + \log p(\boldsymbol{x}_T|\boldsymbol{c},\boldsymbol{z}) + \sum_{t=1}^T \log p_\phi(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t\boldsymbol{c},\boldsymbol{z}) - \log q(\boldsymbol{z}|\boldsymbol{x}_{1:T}) \\
&- \sum_{t=1}^T \log q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{c},\boldsymbol{x}_0)]
\end{aligned}$$

(d) Une factorisation de $q(\boldsymbol{x}_{1:T},\boldsymbol{z}|\boldsymbol{c},\boldsymbol{x}_0)$ :

$$\begin{aligned}
q(\boldsymbol{x}_{1:T},\boldsymbol{z}|\boldsymbol{c},\boldsymbol{x}_0) &= q(\boldsymbol{z}|\boldsymbol{x}_{1:T})\ q(\boldsymbol{x}_{1:T}|\boldsymbol{c},\boldsymbol{x}_0) \\
&= q(\boldsymbol{z}|\boldsymbol{x}_{1:T})\ \prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{c},\boldsymbol{x}_0)
\end{aligned}$$

Donc :

$$\boxed{q(\boldsymbol{x}_{1:T},\boldsymbol{z}|\boldsymbol{c},\boldsymbol{x}_0) = q(\boldsymbol{z}|\boldsymbol{x}_{1:T})\ \prod_{t=1}^T q(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{c},\boldsymbol{x}_0)}$$

(e)

**Question 4** (1-2-1-1-2-3)**.** (**Réseaux antagonistes génératifs**)

Dans cette question, nous souhaitons analyser la dynamique de l'entrainement de GAN sous l'ascension-descente de gradient. On dénote les paramètres du critique et du générateur respectivement par $\psi$ et $\theta$. La fonction objectif considérée est la Jensen-Shannon (standard):

$$\mathcal{L}(\psi,\theta) = \mathbb{E}_{p_D}\log(\sigma(C_\psi(x))) + \mathbb{E}_{p_\theta}\log(\sigma(-C_\psi(x)))$$

où $\sigma$ est la fonction logistique. Pour simplifier l'exposition du problème, nous considérerons le système à temps continu qui résulte du système à temps discret (alternant) quand le taux d'apprentissage, $\eta > 0$, approche zéro:

$$\begin{matrix} \psi^{(k+1)} = \psi^{(k)} + \eta v_\psi(\psi^{(k)},\theta^{(k)}) \\ \theta^{(k+1)} = \theta^{(k)} + \eta v_\theta(\psi^{(k+1)},\theta^{(k)}) \end{matrix} \quad \xRightarrow{\eta\to 0^+} \quad \begin{matrix} \dot\psi = v_\psi(\psi,\theta) \\ \dot\theta = v_\theta(\psi,\theta) \end{matrix} \qquad \begin{matrix} v_\psi(\psi,\theta) \coloneqq \nabla_\psi \mathcal{L}(\psi,\theta) \\ v_\theta(\psi,\theta) \coloneqq -\nabla_\theta \mathcal{L}(\psi,\theta) \end{matrix}$$

Le but est d'étudier la stabilité de l'algorithme d'entrainement. Pour cette raison, nous utiliserons un cadre simple: les données d'entrainement et générées ont le même support sur $\mathbb{R}$. De plus, $p_D = \delta_0$

et $p_\theta = \delta_\theta$. Cela signifie que les deux sont des distributions de Dirac[2] qui sont centrées en $x = 0$, pour les vraies données, et en $x = \theta$, pour les données générées.

Le critique, $C_\psi : \mathbb{R} \to \mathbb{R}$, est $C_\psi(x) = \psi_0 x + \psi_1$.

4.1 Dérivez les expressions pour le champ de "vélocité", $v$, du système dynamique dans l'espace de paramètres $(\psi_0, \psi_1, \theta)$, et trouvez les points stationnaires $(\psi_0^*, \psi_1^*, \theta^*)$.[3]

4.2 Dérivez $J^*$, la jacobienne $(3 \times 3)$ de $v$ au point $(\psi_0^*, \psi_1^*, \theta^*)$.

Pour qu'un système à temps continu soit localement asymptotiquement stable, il suffit que toutes les valeurs propres de $J^*$ aient des parties réelles négatives. Si ce n'est pas le cas, l'étude nécessaire pour conclure est plus complexe et malheureusement, la vitesse de convergence sera au mieux sublinéaire.

4.3 Trouvez les valeurs propres de $J^*$ et discutez de la stabilité locale autour des points stationnaires.

Maintenant, introduisons une pénalité du gradient à la fonction de perte du critique, $\mathcal{R}_1(\psi) = \mathbb{E}_{p_D}\|\nabla_x C_\psi(x)\|^2$. Le système régularisé devient:

$$\begin{aligned} \dot{\psi} &= \bar{v}_\psi(\psi, \theta) \qquad & \bar{v}_\psi(\psi, \theta) &\coloneqq \nabla_\psi \mathcal{L}(\psi, \theta) - \tfrac{\gamma}{2}\nabla_\psi \mathcal{R}_1(\psi) \\ \dot{\theta} &= \bar{v}_\theta(\psi, \theta) \qquad & \bar{v}_\theta(\psi, \theta) &\coloneqq -\nabla_\theta \mathcal{L}(\psi, \theta) \end{aligned}$$

pour $\gamma > 0$. Répétez les étapes 1-2-3 pour ce système modifié et comparez la stabilité des deux systèmes.

4.4 Dérivez les expressions pour le champ de "vélocité", $v$, du système dynamique dans l'espace de paramètres $(\psi_0, \psi_1, \theta)$, et trouvez les points stationnaires $(\psi_0^*, \psi_1^*, \theta^*)$.[4]

4.5 Dérivez $J^*$, la jacobienne $(3 \times 3)$ de $v$ au point $(\psi_0^*, \psi_1^*, \theta^*)$.

4.6 Trouvez les valeurs propres de $J^*$ et discutez de la stabilité locale autour des points stationnaires.

Dans la partie pratique du devoir, vous pourrez vérifier empiriquement vos conclusions.

**Réponse :**

1. Calculons les expressions du champ de vélocity $v$.
   On a :

$$\begin{aligned} \mathcal{L}(\psi, \theta) &= \mathbb{E}_{p_D} \log(\sigma(C_\psi(x))) + \mathbb{E}_{p_\theta} \log(\sigma(-C_\psi(x))) \\ &= \mathbb{E}_{p_D} \log(\sigma(\psi_0 x + \psi_1)) + \mathbb{E}_{p_\theta} \log(\sigma(-\psi_0 x - \psi_1)) \\ &= \mathbb{E}_{p_D} \log(\frac{1}{1 + e^{-\psi_0 x - \psi_1}}) + \mathbb{E}_{p_\theta} \log(\frac{1}{1 + e^{\psi_0 x + \psi_1}}) \\ &= \int_{\mathbb{R}} p_D \log(\frac{1}{1 + e^{-\psi_0 x - \psi_1}}) dx + \int_{\mathbb{R}} p_\theta \log(\frac{1}{1 + e^{\psi_0 x + \psi_1}}) dx \\ &= \int_{\mathbb{R}} \delta_0(x) \log(\frac{1}{1 + e^{-\psi_0 x - \psi_1}}) dx + \int_{\mathbb{R}} \delta_\theta(x) \log(\frac{1}{1 + e^{\psi_0 x + \psi_1}}) dx \\ &= \log(\frac{1}{1 + e^{-\psi_1}}) + \log(\frac{1}{1 + e^{\psi_0 \theta + \psi_1}}) \end{aligned}$$

2. Si $p_X = \delta_z$, alors $p(X = z) = 1$.
3. Pour trouver les points stationnaires, fixez $v = 0$ et résolvez les équations pour chaque paramètre.
4. Pour trouver les points stationnaires, fixez $v = 0$ et résolvez les équations pour chaque paramètre.

Ainsi, on a :

$$\boxed{\mathcal{L}(\psi,\theta) = -\log(1+e^{-\psi_1}) - \log(1+e^{\psi_0\theta+\psi_1})}$$

et les composantes du champs $v$ sont :

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \psi_0} &= -\frac{\theta e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} \\
\frac{\partial \mathcal{L}}{\partial \psi_1} &= \frac{e^{-\psi_1}}{1+e^{-\psi_1}} - \frac{e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} \\
-\frac{\partial \mathcal{L}}{\partial \theta} &= \frac{\psi_0 e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}}
\end{aligned}$$

Pour trouver les points stationnaires $(\psi_0^*, \psi_1^*, \theta^*)$, il suffit de résoudre $v = 0$:

$$v = 0 \iff \frac{\partial \mathcal{L}}{\partial \psi_0} = 0 \text{ et } \frac{\partial \mathcal{L}}{\partial \psi_1} = 0 \text{ et } -\frac{\partial \mathcal{L}}{\partial \theta} = 0$$

On a alors :

$$\begin{aligned}
& \frac{\partial \mathcal{L}}{\partial \psi_0} && = 0 \\
\iff \quad & -\frac{\theta e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} && = 0 \\
\iff \quad & -\theta e^{\psi_0\theta+\psi_1} && = 0 \\
\iff \quad & \theta && = 0
\end{aligned}$$

$$\begin{aligned}
& \frac{\partial \mathcal{L}}{\partial \psi_1} && = 0 \\
\iff \quad & \frac{e^{-\psi_1}}{1+e^{-\psi_1}} - \frac{e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} && = 0
\end{aligned}$$

En utilisant le fait que $\theta^* = 0$, on a :

$$\begin{aligned}
& \frac{\partial \mathcal{L}}{\partial \psi_1} && = 0 \\
\iff \quad & \frac{e^{-\psi_1}}{1+e^{-\psi_1}} - \frac{e^{\psi_1}}{1+e^{\psi_1}} && = 0 \\
\iff \quad & e^{-\psi_1}(1+e^{\psi_1}) - e^{\psi_1}(1+e^{-\psi_1}) && = 0 \\
\iff \quad & e^{-\psi_1} + 1 - e^{\psi_1} - 1 && = 0 \\
\iff \quad & e^{-\psi_1} && = e^{\psi_1} \\
\iff \quad & -\psi_1 && = \psi_1 \\
\iff \quad & \psi_1 = 0
\end{aligned}$$

et enfin,

$$\begin{aligned}
& -\frac{\partial \mathcal{L}}{\partial \theta} && = 0 \\
\iff \quad & \frac{\psi_0 e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} && = 0 \\
\iff \quad & \psi_0 && = 0
\end{aligned}$$

Donc les points stationnaires $(\psi_0^*, \psi_1^*, \theta^*) = (0, 0, 0)$.

2. Calcul de $J^*$, la jacobienne $(3 \times 3)$ de $v$ au point $(\psi_0^*, \psi_1^*, \theta^*)$.
   En posant :

$$v = \begin{pmatrix} \frac{\partial \mathcal{L}}{\partial \psi_0} \\ \frac{\partial \mathcal{L}}{\partial \psi_1} \\ -\frac{\partial \mathcal{L}}{\partial \theta} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

   On a alors l'expression de la Jacobienne $J^*$ aux points stationaires :

$$\begin{aligned}
J^* &= \begin{pmatrix} \frac{\partial v_1}{\psi_0} & \frac{\partial v_1}{\psi_1} & \frac{\partial v_1}{\partial \theta} \\ \frac{\partial v_2}{\psi_0} & \frac{\partial v_2}{\psi_1} & \frac{\partial v_2}{\partial \theta} \\ \frac{\partial v_3}{\psi_0} & \frac{\partial v_3}{\psi_1} & \frac{\partial v_3}{\partial \theta} \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}
\end{aligned}$$

3. Calcul des valeurs propres de $J^*$ :

$$J^* - \lambda I = \begin{pmatrix} -\lambda & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} - \lambda & 0 \\ \frac{1}{2} & 0 & -\lambda \end{pmatrix}$$

   On résout ensuite :

$$\begin{aligned}
\det(J^* - \lambda I) &= 0 \\
\begin{vmatrix} -\lambda & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} - \lambda & 0 \\ \frac{1}{2} & 0 & -\lambda \end{vmatrix} &= 0 \\
-\lambda \begin{vmatrix} -\frac{1}{2} - \lambda & 0 \\ 0 & -\lambda \end{vmatrix} + \frac{1}{2} \begin{vmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} - \lambda & 0 \end{vmatrix} &= 0 \\
-\lambda^3 - \frac{1}{2}\lambda^2 - \frac{1}{4}\lambda - \frac{1}{8} &= 0 \\
(\lambda + \frac{1}{2})(-\lambda^2 - \frac{1}{4}) &= 0
\end{aligned}$$

Les solutions à $(\lambda+\frac{1}{2})(-\lambda^2-\frac{1}{4})$ sont :

$$\lambda_1 = \frac{-1}{2}$$
$$\lambda_2 = \frac{-1}{2}i$$
$$\lambda_3 = \frac{1}{2}i$$

Toutes les valeures propres réelles ne sont pas négatives. On peut ainsi conclure qu'il n'y a pas de stabilité locale autour des points stationaires.

4. On a cette fois-ci :

$$\begin{aligned}
\bar{v}_\psi(\psi,\theta) &= \nabla_\psi \mathcal{L}(\psi,\theta) - \frac{\gamma}{2}\nabla_\psi \mathcal{R}_1(\psi) \\
&= v_\psi(\psi,\theta) - \frac{\gamma}{2}\nabla_\psi \mathcal{R}_1(\psi) \\
&= v_\psi(\psi,\theta) - \frac{\gamma}{2}\nabla_\psi \mathbb{E}_{p_D}\|\nabla_x C_\psi(x)\|^2 \\
&= v_\psi(\psi,\theta) - \frac{\gamma}{2}\nabla_\psi \mathbb{E}_{p_D}\|\psi_0\|^2 \\
&= v_\psi(\psi,\theta) - \frac{\gamma}{2}\nabla_\psi \psi_0^2 \\
&= v_\psi(\psi,\theta) - \frac{\gamma}{2}\begin{pmatrix}2\psi_0\\0\end{pmatrix} \\
&= \begin{pmatrix}\frac{\partial\mathcal{L}}{\partial\psi_0} - \gamma\psi_0 \\ \frac{\partial\mathcal{L}}{\partial\psi_1}\end{pmatrix} \\
&= \begin{pmatrix}-\frac{\theta e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} - \gamma\psi_0 \\ \frac{e^{-\psi_1}}{1+e^{-\psi_1}} - \frac{e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}}\end{pmatrix}
\end{aligned}$$

Les points stationaires $\psi_0^*$ et $\psi_1^*$ restent $\psi_0^*=0$ et $\psi_1^*=0$.
Pour $\theta$, en remplaçant $\psi_0^*=0$ et $\psi_1^*=0$:

$$\begin{aligned}
& -\frac{\theta e^{\psi_0\theta+\psi_1}}{1+e^{\psi_0\theta+\psi_1}} - \gamma\psi_0 && = 0 \\
\iff \quad & -\frac{\theta}{2} && = 0 \\
\iff \quad & \theta && = 0
\end{aligned}$$

Donc les points stationnaires restent les mêmes : $\boxed{(\psi_0^*,\psi_1^*,\theta^*)=(0,0,0)}$.

5. Calcul de la nouvelle matrice jacobienne aux points stationaires :

$$\begin{aligned}
J^* &= \begin{pmatrix} \frac{\partial\bar{v}_1}{\psi_0} & \frac{\partial\bar{v}_1}{\psi_1} & \frac{\partial\bar{v}_1}{\partial\theta} \\ \frac{\partial\bar{v}_2}{\psi_0} & \frac{\partial\bar{v}_2}{\psi_1} & \frac{\partial\bar{v}_2}{\partial\theta} \\ \frac{\partial\bar{v}_3}{\psi_0} & \frac{\partial\bar{v}_3}{\psi_1} & \frac{\partial\bar{v}_3}{\partial\theta} \end{pmatrix} \\
&= \begin{pmatrix} -\gamma & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}
\end{aligned}$$

6. Calcul des valeurs propres de $J^*$ :

$$\det(J^* - \lambda I) = 0$$

$$\begin{vmatrix} -\gamma - \lambda & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} - \lambda & 0 \\ \frac{1}{2} & 0 & -\lambda \end{vmatrix} = 0$$

$$-\gamma - \lambda \begin{vmatrix} -\frac{1}{2} - \lambda & 0 \\ 0 & -\lambda \end{vmatrix} + \frac{1}{2} \begin{vmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} - \lambda & 0 \end{vmatrix} = 0$$

$$-\lambda^3 - (\gamma + \frac{1}{2})\lambda^2 - (\frac{1}{2}\gamma + \frac{1}{4})\lambda - \frac{1}{8} = 0$$

En remarquant que

$$-\lambda^3 - (\gamma + \frac{1}{2})\lambda^2 - (\frac{1}{2}\gamma + \frac{1}{4})\lambda - \frac{1}{8} \quad = \quad (\lambda + \frac{1}{2})(-\lambda^2 - \gamma\lambda - \frac{1}{4})$$

On a alors la première valeur propre $\lambda_1 = -\frac{1}{2}$.
Interessons nous à l'équation

$$-\lambda^2 - \gamma\lambda - \frac{1}{4} = 0$$

Le discriminant $\Delta$ de cette equation du second degré est égale à :

$$\Delta = \gamma^2 - 1$$

Si $\gamma < 1$, les valeurs propres $\lambda_2$ et $\lambda_3$ sont complexes. Calculons les dans ce cas-ci :

$$\lambda_2 = \frac{\gamma - i\sqrt{1-\gamma^2}}{-2} = -\frac{\gamma}{2} + i\frac{\sqrt{1-\gamma^2}}{2}$$

$$\lambda_3 = -\frac{\gamma}{2} - i\frac{\sqrt{1-\gamma^2}}{2}$$

Sachant que $\gamma > 0$, alors dans le cas où $\gamma < 1$, on a les parties réelles de toutes les valeurs propres qui sont négatives.
Regardons ce qu'il en est lorsque $\gamma \geq 1$ :
Dans ce cas on a :

$$\lambda_2 = \frac{\gamma - \sqrt{1-\gamma^2}}{-2} = -\frac{\gamma}{2} + \frac{\sqrt{1-\gamma^2}}{2}$$

$$\lambda_3 = -\frac{\gamma}{2} - \frac{\sqrt{1-\gamma^2}}{2}$$

Sachant que $\gamma \geq 1$, et que $\gamma > \sqrt{1-\gamma^2}$, on a bien $\lambda_2 < 0$. De plus $\lambda_3 < 0$.
En effet :

$$\gamma^2 > \gamma^2 - 1$$

$$\sqrt{\gamma^2} > \sqrt{\gamma^2 - 1}$$

$$\gamma > \sqrt{\gamma^2 - 1}$$

Récapitulatif :
On a, $\forall \gamma > 0$, $\lambda_1 < 0$, $\lambda_2 < 0$ et $\lambda_3 < 0$.

Conclusion :
On a ainsi montré que $\forall \gamma > 0$, toutes les valeurs propres possèdent une partie réelle strictement négative. On peut ainsi conclure qu'**il y a ici stabilité locale autour des points stationaires**.

**Question 5** (20 pts)**.** (**Revue de papier**)

dans cette question, vous allez écrire **une page** révision de Apprentissage auto-supervisé. Veuillez structurer votre examen dans les sections suivantes:

1. **Sommaire [5 pts]**:
   (a) De quoi parle cet article ?
   (b) Quelle est la principale contribution ?
   (c) Décrivez l'approche principale et les résultats. Seulement des faits, pas encore d'opinions.
2. **Forces [5 pts]**:
   (a) Y a-t-il un nouvel aperçu théorique ?
   (b) Ou un progrès empirique important ? Ont-ils résolu un problème permanent ?
   (c) Ou une bonne formulation pour un nouveau problème ?
   (d) Des résultats concrets (code, algorithme, etc.) ?
   (e) Les expériences sont-elles bien exécutées ?
   (f) Utile pour la communauté en général ?
3. **Faiblesses [5 pts]** :
   (a) Que peut-on faire de mieux ?
   (b) Des bases de référence manquantes ? Des ensembles de données manquants ?
   (c) Des choix de conception bizarres dans l'algorithme ne sont-ils pas bien expliqués ? Qualité de l'écriture ?
   (d) Est-ce qu'il y a suffisamment de nouveauté dans ce qu'ils proposent ? Variation mineure de travaux antérieurs ?
   (e) Pourquoi devrait-on s'en soucier ? Le problème est-il intéressant et important ?
4. **Reflets [5 pts]**:
   (a) Quel est le lien avec d'autres concepts que vous avez vus dans la classe ?
   (b) Quelles sont les prochaines orientations de recherche dans ce domaine ?
   (c) Quelles nouvelles idées (directement ou indirectement liées) ce document vous a-t-il donné ? Qu'est-ce que tu voudrais essayer ?

**Réponse :**
**Résumé**:
Cet article présente SimCLR (SIMple framework for Contrastive Learning of visual Representations). SimCLR est un algorithme d'apprentissage auto-supervisé contrastif simplifié, sans nécessiter d'architectures spécialisées ou de banque de mémoire. Les contributions apportées sont : (1) la

composition des augmentations de données joue un rôle crucial, (2) une transformation non linéaire apprenable entre la représentation et la perte contrastive améliore sensiblement la qualité de la représentation, et enfin (3) l'apprentissage contrastif gagne en performance avec de plus grandes tailles de lots et de plus d'étapes d'entraînement comparé à l'apprentissage supervisé.
Les résultats obtenus permettent de surpasser considérablement les méthodes précédentes d'apprentissage auto-supervisé et semi-supervisé sur ImageNet. Un classificateur linéaire formé sur des représentations auto-supervisées apprises par Sim-CLR atteint une précision 76,5% top-1, ce qui représente une amélioration relative de 7% par rapport à ce qui se faisait précedemment, correspondant aux performances d'un ResNet-50 supervisé. De plus, lorsqu'on procède à un ajustement (fine-tune) sur 1% des labels, une précision de 85,8% top-5 est obtenue, surpassant AlexNet avec 100 fois moins de labels.
**Forces**:
Je trouve que la contribution clé de l'article est **l'utilisation d'augmentations de données** : SimCLR crée des paires d'images pour en apprendre la similarité. Si nous saisissions la même image deux fois, il n'y aurait aucun effet d'apprentissage. Par conséquent, chaque paire d'images est créée en appliquant des augmentations ou des transformations à l'image.
Différentes augmentations sont appliquées telles que le redimensionnement, la distorsion des couleurs, le flou, le bruit et bien plus encore. Différentes parties de l'image sont également sélectionnées, ce qui est important pour que le modèle apprenne une représentation cohérente. L'image peut être recadrée dans une vue globale et locale (image complète et partie recadrée de l'image) ou des vues adjacentes peuvent être utilisées (deux recadrages de différentes parties de l'image). Il est important de préciser que deux images augmentées contiennent le même objet.
Ensuite, ces paires sont passées dans un réseau neuronal convolutif (ResNet) pour créer une représentation de caractéristiques pour chacune des images. Une fois que la représentation vectorielle d'une image d'entrée est calculée, cette sortie est transmise à un MLP avec une couche cachée. Ce MLP n'est utilisé que pendant l'entraînement et pour affiner la représentation des caractéristiques des images d'entrée.
Ce que je trouve très intéréssant est aussi **la manière dont est évalué la performance du modèle**. En effet, le modèle essaye de minimiser la distance entre les images contenant le même objet et de maximiser la distance entre les images contenant des objets très différents (constrastive learning). Une contribution majeure de l'article SimCLR est **l'utilisation de la fonction perte NT-Xent**. NT-Xent (normalized temperature-scaled cross entropy loss) est une fonction de perte qui a une propriété particulièrement utile : différents exemples sont pondérés efficacement, ce qui permet au modèle d'apprendre beaucoup plus efficacement à partir de représentations vectorielles éloignées les unes des autres même si leur origine est la même image. De plus, cette fonction permet de retrouver une certaine attraction chez les images similaires, c'est-à-dire que des images similaires sont apprises à être représentées plus près les unes des autres.
Au de-là des nouveaux aspects techniques précedemment cités que j'ai trouvé pertinents et innovants, le papier est très bien écrit, les travaux de recherche sont tous clairement énoncés et résumés. Les expériences semblent être bien exécutées, l'algorithme utilisé est bien retranscrit et expliqué. L'évaluation est approfondie et l'analyse est bonne. Les résultats sont intéressants pour la communeauté : ils montrent que SimCLR surpasse toutes les autres méthodes auto-supervisées. De plus, le modèle performe également bien en battant la méthode d'entraînement supervisée sur de nombreux datasets. Les discussions supplémentaires en annexe sont également utiles pour comprendre le modèle.

**Faiblesse**:
En réalité, le papier ne présente pas de nouveauté technique significative. En effet, plusieurs idées étaient déjà existantes dans des papiers antérieurs, tel que l'idée de faire concorder les représentations d'une image par petites transformations par exemple. De plus, l'approche 'Contrastive visual representation learning', approche qui permet d'apprendre les représentations en opposant les 'positive pairs' aux 'negative pairs' n'est pas une nouveauté à proprement parler. Presque tous les composants du papier sont apparus dans des travaux antérieurs ou présentent des variations mineurs, même si une comparaison des choix de conception avec ceux des travaux antérieurs est fournie en annexe.
Enfin, je trouve cela dommage de ne pas avoir testé différentes architectures de CNN, et seulement avoir utilisé l'architecture ResNet.

**Reflexion**:
Le papier traite un sujet qu'on a pu évoquer en cours : "Self-Supervised Learning". En effet, nous avions pu voir une approche du "Contrastive learning" où le but était de minimiser la variance intra-image à travers différentes augmentations, et maximiser la variance across-image. De plus, nous avions introduit les "Negative Pairs" et "Positive Pairs" évoqués dans le papier, ainsi que les techniques visant à transformer les images (la recherche de 'contexte' et de position relative, les rotations, etc...).
De futurs orientations de recherche seraient d'essayer d'optimiser le modèle. En effet, il serait interessant à mon avis de tester différents architectures de CNN, en plus de ResNet, pour constater si oui ou non, on parvient à obtenir de meilleurs résultats.