DM modèles génératifs pour l'image

Nissim Maruani

March 3, 2022

Question 1

Chaque coordonnée de la variable aléatoire $K = k * W$ est une loi normale car c'est la somme de $M \times N$ lois normales. K est un champ gaussien de moyenne:

$$\begin{aligned}
\mathbb{E}(K(x)) &= \mathbb{E}((k * W)(x)) && (1)\\
&= \mathbb{E}(\sum_y k(y)W(x-y)) && (2)\\
&= \sum_y \mathbb{E}(k(y)W(x-y)) && (3)
\end{aligned}$$

Puisque l'image k est constante (W est la variable aléatoire):

$$\begin{aligned}
&= \sum_y k(y)\mathbb{E}(W(x-y)) && (4)\\
&= 0 && (5)
\end{aligned}$$

Cette moyenne est bien nulle. La covariance vaut:

$$\begin{aligned}
Cov(K(x), K(y)) &= \mathbb{E}((K(x) - 0)(K(y) - 0)) && (6)\\
&= \mathbb{E}((\sum_z k(z)W(x-z))(\sum_z k(z)W(y-z))) && (7)\\
&= \mathbb{E}((\sum_z k(x-z)W(z))(\sum_z k(y-z)W(z))) && (8)
\end{aligned}$$

Puisque les lois normales $W(z)$ sont indépendantes, centrées et de variance 1:

$$\begin{aligned}
&= \sum_z k(x-z)k(y-z) && (9)\\
&= \sum_z k(z)k(y-x+z) && (10)\\
&= \sum_z k(z)\tilde{k}(x-y-z) && (11)\\
&= k * \tilde{k}(x-y) && (12)
\end{aligned}$$

Question 2

Soient k et h deux images ayant le même module de Fourier. D'après l'énoncé, on a pour tout ξ: $\widehat{k * \tilde{k}}(\xi) = \widehat{h * \tilde{h}}(\xi)$. La transformée de Fourier étant un isomorphisme, on a pour tout x: $k * \tilde{k}(x) = h * \tilde{h}(x)$, on reconnait la covariance de la question précédente.

Les deux champs gaussiens $k * W$ et $h * W$ ont tous les deux une moyenne nulle et la même covariance: ils sont égaux, donc $SNDA(k) = SNDA(h)$.

Question 3

Si $t_k = t_h$, alors $\widehat{t_k} = \widehat{t_h}$, ce qui signifie que k et h ont le même module de Fourier. D'après la Question 2, $SNDA(k) = SNDA(h)$, K et H ont la même distribution.

Réciproquement, si K et H ont la même distribution, on a $Cov(K) = Cov(H)$ et donc $k * \tilde{k} = h * \tilde{h}$ d'après la Question 1. D'après l'énoncé, ceci équivaut à $MN|\widehat{k}(\xi)|^2 = MN|\widehat{h}(\xi)|^2$. Les modules de Fourier sont égaux, donc $t_k = t_h$. On a bien équivalence.

Question 4

On a d'après l'énoncé:

$$\begin{aligned}
\mathbb{E}(\|K - H\|^2) &= \mathbb{E}(MN\|\widehat{K} - \widehat{H}\|^2) && (13)\\
&= MN\mathbb{E}(< \widehat{K} - \widehat{H}, \widehat{K} - \widehat{H} >) && (14)\\
&= MN(\mathbb{E}(|\widehat{K}|^2) + \mathbb{E}(|\widehat{H}|^2) - 2Re(< \widehat{K}, \widehat{H} >)) && (15)\\
&= MN\sum_{\xi} \mathbb{E}(|\widehat{K}(\xi)|^2) + \mathbb{E}(|\widehat{H}(\xi)|^2) - 2\mathbb{E}(Re(\widehat{K}(\xi)\overline{\widehat{H}(\xi)})) && (16)
\end{aligned}$$

Question 5

Si W est un bruit blanc gaussien standard, alors :

$$\mathbb{E}(|\widehat{W}(\xi)|^2) = \mathbb{E}(|\frac{1}{MN}\sum_{x} W(x)\omega_1^{-x_1\xi_1}\omega_2^{-x_2\xi_2}|^2) \qquad (17)$$

Les $W(x)$ sont indépendants et de moyenne nulle :

$$\begin{aligned}
&= \frac{1}{(MN)^2}\sum_{x} \mathbb{E}((W(x)\omega_1^{-x_1\xi_1}\omega_2^{-x_2\xi_2})\overline{(W(x)\omega_1^{-x_1\xi_1}\omega_2^{-x_2\xi_2})}) && (18)\\
&= \frac{1}{(MN)^2}\sum_{x} 1 && (19)\\
&= \frac{1}{MN} && (20)
\end{aligned}$$

Donc $\mathbb{E}(|\widehat{K}(\xi)|^2) = \mathbb{E}(|\widehat{k * W}(\xi)|^2) = \mathbb{E}((MN)^2|\widehat{k}(\xi)|^2|\widehat{W}(\xi)|^2) = MN|\widehat{k}(\xi)|^2$, et de la même manière $\mathbb{E}(|\widehat{H}(\xi)|^2) = MN|\widehat{h}(\xi)|^2$.

Question 6

On a par Cauchy-Schwarz $< \widehat{K}(\xi), \widehat{H}(\xi) > \leq |\widehat{K}(\xi)||\widehat{H}(\xi)|$. Donc $\mathbb{E}(Re(\widehat{K}(\xi)\overline{\widehat{H}(\xi)})) \leq \mathbb{E}(|\widehat{K}(\xi)||\widehat{H}(\xi)|)$.

De plus, $\mathbb{E}(|\widehat{K}(\xi)||\widehat{H}(\xi)|) = \mathbb{E}(MN|\widehat{k}(\xi)||\widehat{W}(\xi)|MN|\widehat{h}(\xi)||\widehat{W}(\xi)|) = MN|\widehat{k}(\xi)||\widehat{h}(\xi)|$. On a bien:

$$\mathbb{E}(Re(\widehat{K}(\xi)\overline{\widehat{H}(\xi)})) \leq \mathbb{E}(|\widehat{K}(\xi)||\widehat{H}(\xi)|) \leq MN|\widehat{k}(\xi)||\widehat{h}(\xi)|$$

Question 7

D'après les questions précédentes, on a pour tout (K, H):

$$\begin{aligned}
\mathbb{E}(\|K-H\|^2) &\geq MN\sum_{\xi} MN|\widehat{k}(\xi)|^2 + MN|\widehat{h}(\xi)|^2 - 2MN|\widehat{k}(\xi)||\widehat{h}(\xi)| && (21)\\
&= (MN)^2 \sum_{\xi}(|\widehat{k}(\xi)| - |\widehat{h}(\xi)|)^2 && (22)\\
&= (MN)^2\|\widehat{t}_k - \widehat{t}_h\|^2 && (23)\\
&= MN\|t_k - t_h\|^2 && (24)
\end{aligned}$$

En passant à l'inf, $d_w(SNDA(k), SNDA(h))^2 \geq MN\|t_k - t_h\|^2$

Question 8

Si on prend $(K, H) = (t_k * W, t_h * W)$, on a $\widehat{K} \in \mathbf{R}$ et $\widehat{H} \in \mathbf{R}$, ce qui donne l'égalité dans Cauchy-Schwarz et donc $\mathbb{E}(\|K-H\|^2) = MN\|t_k - t_h\|^2$.

Autrement dit, la distance de Wasserstein entre les SNDA de noyaux k et h est donnée en forme close par la norme 2 (moyennant une constante MN) entre les deux textons t_k et t_h:

$$d_w(SNDA(k), SNDA(h))^2 = MN\|t_k - t_h\|^2$$

Question 9

D'après la question précédente, on a:

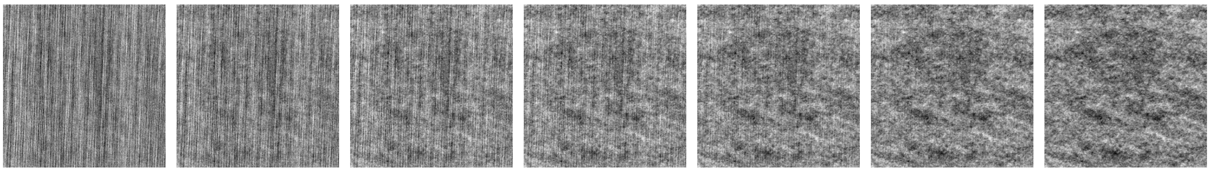
$$\min_{\mu \sim SNDA(k)} (1-\rho)d_w(\mu_0, \mu) + \rho d_w(\mu_1, \mu) = \min_{t_k}(1-\rho)MN\|t_{k_0} - t_k\|^2 + \rho MN\|t_{k_1} - t_k\|^2$$

L'unique minimum de ce problème convexe est atteint quand le gradient s'annule, i.e. quand: $MN(2t_p - 2((1-\rho)t_{k_0} + \rho t_{k_1}) = 0 \Leftrightarrow t_p = (1-\rho)t_{k_0} + \rho t_{k_1}$

Le texton t_p de la distribution μ_p réalisant le minimum est le barycentre euclidien des deux textons.

Question 10

Voici ce qu'on obtient en faisant varier ρ de 0 à 1 sur 7 valeurs pour le couple d'images *mur.png* et *tissu.png* (ne pas hésiter à zoomer sur le PDF):



(a) $\rho = 0$ (b) $\rho = \frac{1}{6}$ (c) $\rho = \frac{2}{6}$ (d) $\rho = \frac{3}{6}$ (e) $\rho = \frac{4}{6}$ (f) $\rho = \frac{5}{6}$ (g) $\rho = 1$

Figure 1: From *tissu.png* to *wall.png*

Question 11

La méthode de Gatys peut tout à fait s'adapter au synthèse de barycentre de texture en choisissant pour loss $(1-\rho)E_w(x, u_0) + \rho E_w(x, u_1)$. Voici les résultats:

(a) $\rho = 0$ (b) $\rho = \frac{1}{6}$ (c) $\rho = \frac{2}{6}$ (d) $\rho = \frac{3}{6}$ (e) $\rho = \frac{4}{6}$ (f) $\rho = \frac{5}{6}$ (g) $\rho = 1$

Figure 2: From *wood.png* to *wall.png*

(a) $\rho = 0$ (b) $\rho = \frac{1}{6}$ (c) $\rho = \frac{2}{6}$ (d) $\rho = \frac{3}{6}$ (e) $\rho = \frac{4}{6}$ (f) $\rho = \frac{5}{6}$ (g) $\rho = 1$

Figure 3: From *briques.png* to *pebbles.png*

Le résultat est particulièrement convaincant: sur le dernier exemple, on a vraiment l'impression que les briques se transforment en galets.