Image Denoising : Course 1

Maxence Gollier

October 2023

1 Exercices

1.1 Exercise 4.1

Let $\lambda > 0$, and $X \sim \text{Pois}(\lambda)$

- $\mathbb{E}[X] = \sum_{k\geq 0} kp_k = e^{-\lambda}\sum_{k\geq 0} k\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k\geq 1}\frac{\lambda^{k-1}}{(k-1)!} = \lambda$
- $\mathbb{V}[X] = \sum_{k\geq 0} k^2 p_k - \lambda^2 = e^{-\lambda}\sum_{k\geq 0} k^2\frac{\lambda^k}{k!} - \lambda^2 = \lambda e^{-\lambda}\sum_{k\geq 1} k\frac{\lambda^{k-1}}{(k-1)!} - \lambda^2 =$
  $\lambda e^{-\lambda}(\sum_{k\geq 1}(k-1)\frac{\lambda^{k-1}}{(k-1)!} + \sum_{k\geq 1}\frac{\lambda^{k-1}}{(k-1)!}) - \lambda^2 = \lambda e^{-\lambda}(\lambda e^{\lambda} + e^{\lambda}) - \lambda^2 = \lambda.$

1.2 Exercise 4.2

Let $n > 0$, $\lambda_i > 0$, $X_i \sim \text{Pois}(\lambda_i)$ be n independant Poisson variables. Then, $Y = \sum_i X_i \sim \text{Pois}(\sum_i \lambda_i)$.

proof We only have to prove for the case $n = 2$, as the other cases can then be deduced by induction. We have

$$\begin{aligned}
\mathbb{P}[X_1 + X_2 = k] &= \sum_{l}^{k} \mathbb{P}[X_1 = l \cap X_2 = k - l] \\
&= \sum_{l}^{k} \mathbb{P}[X_1 = l]\mathbb{P}[X_2 = k - l] \\
&= e^{-(\lambda_1+\lambda_2)}\frac{1}{k!}\sum_{l}^{k}\lambda_1^l\lambda_2^{k-l}\frac{k!}{l!(k-l)!} \\
&= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^k}{k!}.
\end{aligned}$$

Hence, $Y = X_1 + X_2 \sim \text{Pois}(\lambda)$.

1.3 Exercise 4.3

In the case of a linear variance model, we have $\sigma^2(u) = u$. Hence, $g(u) = \sigma(u) = \sqrt{u}$ which yields

$$f(u) = \int_0^u \frac{cdt}{\sqrt{t}} = 2c\sqrt{u}. \tag{1}$$

1.4 Exercise 4.5

Using the notations from the lecture notes, and using the facts that

- $\langle \tilde{U}, G_i \rangle = \langle U, G_i \rangle + \langle N, G_i \rangle$,
- $\langle N, G_i \rangle = 0$,
- $|\langle N, G_i \rangle|^2 = \sigma^2$.

We get that for any $D\tilde{U} = \sum_{i=1}^M a_i \langle \tilde{U}, G_i \rangle G_i$

$$\begin{aligned}
\mathbb{E}[\|U - D\tilde{U}\|^2] &= \|U\|^2 + \mathbb{E}[\|D\tilde{U}\|^2] - 2\mathbb{E}[\langle U, D\tilde{U} \rangle] \\
&= \|U\|^2 + \sum_{i=1}^M a_i^2(|\langle U, G_i \rangle|^2 + \sigma^2) - 2\sum_{i=1}^M a_i |\langle U, G_i \rangle|^2.
\end{aligned}$$

Taking the gradient with respect to $\mathbf{a}$ we get

$$\nabla_{a_i} \mathbb{E}[\|U - D\tilde{U}\|^2] = 2a_i(|\langle U, G_i \rangle|^2 + \sigma^2) - 2|\langle U, G_i \rangle|^2 = 0.$$

Then, solving for a_i, we get

$$a_i^* = \frac{|\langle U, G_i \rangle|^2}{|\langle U, G_i \rangle|^2 + \sigma^2}.$$

Clearly, $\nabla_a^2 \mathbb{E}[\|U - D\tilde{U}\|^2]$ is diagonal with positive entries on its diagonal which implies that the hessian is positive definite hence we found a minimum. The last part of the theorem is found by substituting a_i^* in $\mathbb{E}[\|U - D\tilde{U}\|^2]$.

1.5 Exercise 4.6

From Theorem 4.2, we get

$$\mathbb{E}[\|U - D_{\text{inf}}\tilde{U}\|^2] = \sum_{i=1}^M \frac{|\langle U, G_i \rangle|^2 \sigma^2}{|\langle U, G_i \rangle|^2 + \sigma^2}.$$

We always have for any $c \geq 1$, for any $i \in \{1, ..., M\}$

$$\begin{aligned}
\frac{|\langle U, G_i \rangle|^2 \sigma^2}{|\langle U, G_i \rangle|^2 + \sigma^2} &\leq |\langle U, G_i \rangle|^2 \\
\frac{|\langle U, G_i \rangle|^2 \sigma^2}{|\langle U, G_i \rangle|^2 + \sigma^2} &\leq \sigma^2 \leq c\sigma^2.
\end{aligned}$$

Hence,

$$\mathbb{E}[\|U - D_{\inf}\tilde{U}\|^2] \leq \sum_{i=1}^{M} \min\{|\langle U, G_i\rangle|^2, c\sigma^2\}.$$

1.6 Exercise 4.7

To show, this it is sufficient to prove that the DCT is orthogonal and that the IDCT is the transpose of the DCT.

Let $j_1, j_2 \in \{0, ..., N-1\}$

- Suppose $j_1 \neq j_2$. Then,

$$\begin{aligned}\langle DCT_{j_1}, DCT_{j_2}\rangle &= 4\sum_{k=0}^{N-1} \alpha_k^2 \cos(\frac{\pi k}{N}(j_1 + \frac{1}{2}))\cos(\frac{\pi k}{N}(j_2 + \frac{1}{2}))\\ &= -\frac{1}{N} + \frac{1}{N}\sum_{k=0}^{N-1}\cos(\frac{\pi k}{N}(j_1 + j_2 + 1)) + \cos(\frac{\pi k}{N}(j_1 - j_2))\\ &= -\frac{1}{N} + \frac{1}{N}\mathcal{R}(\frac{1 - e^{i\pi(j_1+j_2+1)}}{1 - e^{\frac{i\pi}{N}(j_1+j_2+1)}}) + \mathcal{R}(\frac{1 - e^{i\pi(j_1-j_2)}}{1 - e^{\frac{i\pi}{N}(j_1-j_2)}})\end{aligned}$$

  If we suppose j_1 is even and j_2 is odd or vice-versa then,

$$\langle DCT_{j_1}, DCT_{j_2}\rangle = -\frac{1}{N} + \frac{1}{N}\mathcal{R}(\frac{2}{1 - e^{\frac{i\pi}{N}(j_1-j_2)}}) = 0.$$

  If we suppose both j_1 and j_2 are either even or odd then

$$\langle DCT_{j_1}, DCT_{j_2}\rangle = -\frac{1}{N} + \frac{1}{N}\mathcal{R}(\frac{2}{1 - e^{\frac{i\pi}{N}(j_1+j_2+1)}}) = 0.$$

- Suppose $j_1 = j_2 = j$. Then, using the same reasoning as for the last item,

$$\begin{aligned}\langle DCT_j, DCT_j\rangle &= -\frac{1}{N} + \frac{1}{N}\sum_{k=0}^{N-1}\cos(\frac{\pi k}{N}(2j+1)) + \cos(\frac{\pi k}{N}(0))\\ &= 1 - \frac{1}{N} + \frac{1}{N}\mathcal{R}(\frac{2}{1 - e^{\frac{i\pi}{N}(2j+1)}}) = 1.\end{aligned}$$

Therefore, the DCT is orthogonal and it is an isometry in $\mathcal{R}^N$. Its inverse is then given by its transpose which is

$$\begin{aligned}X_j &= \sum_{k=0}^{N-1} 2\alpha_k Y_k \cos(\frac{\pi k}{N}(j + \frac{1}{2}))\\ &= 2\alpha_0 Y_0 + \sum_{k=1}^{N-1} 2\alpha_k Y_k \cos(\frac{\pi k}{N}(j + \frac{1}{2}))\\ &= \beta_0 Y_0 + \sum_{k=1}^{N-1} 2\beta_k Y_k \cos(\frac{\pi k}{N}(j + \frac{1}{2})) = IDCT(Y).\end{aligned}$$

1.7 Exercise 4.8

The constrained problem being convex with linear equality constraints, Slater's conditions for example justify the existence of a point where the Lagrangian of the constrained problem is 0 which is

$$2\sigma_k^2\alpha_k - \lambda = 0$$

2 Experimental Report

2.1 Ponomarenko algorithm

The Ponomarenko algorithm has some similarities with the DCT denoising seen in the course. It takes a noisy image $\tilde{U}$ Extracts $w \times w$ sized patches W_m from it and applies a DCT transform to each one of the patches yielding D_m. The key difference is that it does not require a priori knowledge of the noise level σ which is not known in practice. In order to estimate the variance of the model it $\hat{\sigma}^2$ it proceeds in four steps :

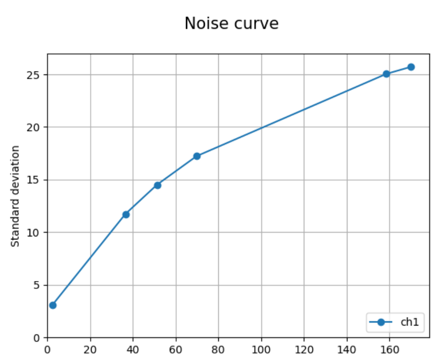
- Subdivide each patches between the low-frequency and the medium/high-frequency parts. This is done using a threshold defined by the user.
- Estimate the block empirical variances for the low-frequencies only.
- Use the low frequencies blocks to estimate the high frequency noise.
- Take a median among the high-frequency estimated variances.

2.2 Experiments

In order to test the method, we check how accurately the noise curve fits the variance noise model which is added by hand.

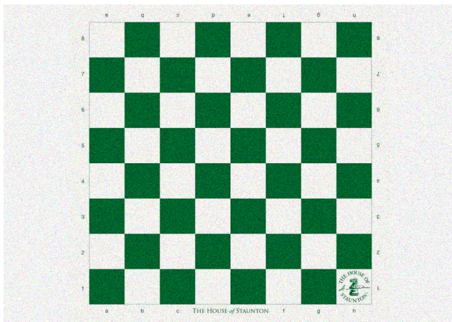
First, we use *building1* and add a noise with variance $\sigma^2(u) = 4u$.

Figure 1: noisy building1 - $\sigma^2(u) = 4u$



We see on Figure 1 that we indeed have a standard deviation proportional to $\sqrt{u}$.

Since the model assumes high frequencies are noise, a highly textured image might harm the algorithm. We test this with a chessboard with the same previous noise.



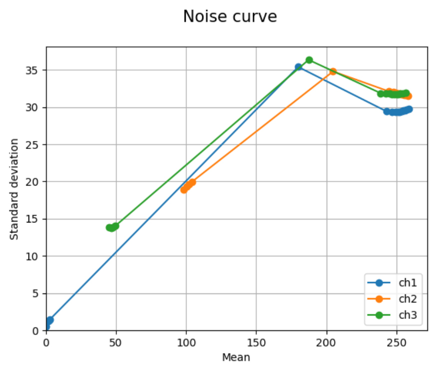


Figure 2: noisy chessboard - $\sigma^2(u) = 4u$

Clearly now, we see on Figure 5 that the algorithm does not behave well in that case.

2.3 DCT Denoising

We use a single image and comment on the algorithm from it.

Figure 3: building1 - original - $\sigma = 30$

We see that the noise does indeed disapear. However, since we use the threshold 3σ, we see that some high frequency details such as the leafs from the

Figure 4: building1 - noisy - $\sigma = 30$

Figure 5: building1 - denoised - $\sigma = 30$

trees or the details on the right wall disapear. This is a compromise : Removing all high frequencies deletes the noise but also deletes the textures from the image.