Lucas Versini
lucas.versini01@gmail.com
lucas.versini@polytechnique.edu

Image Denoising

Homework n°1 - Exercises

Exercise 4.1

Let $X \sim \mathcal{P}(\lambda)$. Then we have:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{+\infty} k\mathbb{P}(X=k) \\
&= e^{-\lambda} \sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=1}^{+\infty} \lambda^k \frac{\cancel{k}}{\cancel{k} \times (k-1)!} \\
&= \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&\overset{l=k-1}{=} \lambda e^{-\lambda} \underbrace{\sum_{l=0}^{+\infty} \frac{\lambda^l}{l!}}_{=e^{\lambda}} \\
&= \lambda
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{k=0}^{+\infty} k^2\mathbb{P}(X=k) \\
&= e^{-\lambda} \sum_{k=0}^{+\infty} k^2 \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=2}^{+\infty} k(k-1) \frac{\lambda^k}{k!} + e^{-\lambda} \sum_{k=1}^{+\infty} k \frac{\lambda^k}{k!} \\
&= \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&\overset{\substack{l=k-2 \\ m=k-1}}{=} \lambda^2 e^{-\lambda} \underbrace{\sum_{l=0}^{+\infty} \frac{\lambda^l}{l!}}_{=e^{\lambda}} + \lambda e^{-\lambda} \underbrace{\sum_{m=0}^{+\infty} \frac{\lambda^m}{m!}}_{=e^{\lambda}} \\
&= \lambda^2 + \lambda
\end{aligned}$$

And $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$.

So we have shown $\boxed{\mathbb{E}[X] = \mathrm{Var}(X) = \lambda}$.

Exercise 4.2

We first prove the result for $n = 2$.
Let $X_1 \sim \mathcal{P}(\lambda_1)$ and $X_2 \sim \mathcal{P}(\lambda_2)$ be independent random variables.
$X_1 + X_2$ is valued in $\mathbb{N}$, and for $k \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = k) &= \mathbb{P}\left(\bigcup_{l=0}^{k} \{X_1 = l, X_2 = k - l\}\right) \text{ by the total probability formula} \\
&= \sum_{l=0}^{k} \mathbb{P}(X_1 = l, X_2 = k - l) \text{ because the union is disjoint} \\
&= \sum_{l=0}^{k} \mathbb{P}(X_1 = l)\mathbb{P}(X_2 = k - l) \text{ because } X_1 \perp\!\!\!\perp X_2 \\
&= \sum_{l=0}^{k} e^{-\lambda_1} \frac{\lambda_1^l}{l!} e^{-\lambda_2} \frac{\lambda_2^{k-l}}{(k-l)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{l=0}^{k} \binom{k}{l} \lambda_1^l \lambda_2^{k-l} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k \text{ by the binomial theorem.}
\end{aligned}$$

This shows that $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$. Then, we prove the result for any $n \geq 1$ by induction.

- The result is clear for $n = 1$.
- Assume the result holds for some $n \geq 1$, and write $X_1 + \cdots + X_{n+1} = (X_1 + \cdots + X_n) + X_{n+1}$.

  By the induction hypothesis, $X_1 + \cdots + X_n \sim \mathcal{P}(\sum_{k=1}^{n} \lambda_k)$.

  We also have that $X_1 + \cdots + X_n$ and X_{n+1} are independent (because X_{n+1} is independent from $X_1, \ldots, X_n$).

  So by what we showed previously, we have $X_1 + \cdots + X_{n+1} \sim \mathcal{P}(\sum_{k=1}^{n} \lambda_k + \lambda_{n+1}) = \mathcal{P}(\sum_{k=1}^{n+1} \lambda_k)$.

This proves the induction, and concludes the exercise.

Exercise 4.3

By Taylor's expansion, for f smooth we have $f(\tilde{u}) \approx f(u) + f'(u)(\tilde{u} - u) \approx f(u) + f'(u)g(u)n$.

To have a variance independent of u, we want $f'(u)g(u)$ to be constant, that is, $f'(u) = \dfrac{c_1}{g(u)}$.
Here, $g(u) = \sqrt{u}$, so $f'(u) = \dfrac{c_1}{\sqrt{u}}$ and $f(u) = 2c_1\sqrt{u} + c_2$.
Setting $c_2 = 0$ and $c_1 = c$ yields $\boxed{f(u) = 2c\sqrt{u}}$.
In that case, $f(\tilde{u}) \approx 2c\sqrt{u} + cn$.

Exercise 4.5

Set $\mathbf{D} = \sum_{i=1}^{M} a(i) G_i G_i^T$.

Since $(G_i)_{i=1,\dots,M}$ is an orthonormal basis, we have $U = \sum_{i=1}^{M} \langle U, G_i \rangle G_i$, and since $\tilde{U} = U + N$, we have $\mathbf{D}\tilde{U} = \sum_{i=1}^{M} a(i) \left(\langle U, G_i \rangle + \langle N, G_i \rangle \right) G_i$. So:

$$
\begin{aligned}
\mathbb{E}\left[\|U - \mathbf{D}\tilde{U}\|^2 \right] &= \mathbb{E}\left[\left\| \sum_{i=1}^{M} \left(\langle U, G_i \rangle - a(i) \left(\langle U, G_i \rangle + \langle N, G_i \rangle \right) \right) G_i \right\|^2 \right] \\
&= \mathbb{E}\left[\left\| \sum_{i=1}^{M} \left((1 - a(i)) \langle U, G_i \rangle - a(i) \langle N, G_i \rangle \right) G_i \right\|^2 \right] \\
&= \mathbb{E}\left[\sum_{i=1}^{M} \left((1 - a(i)) \langle U, G_i \rangle - a(i) \langle N, G_i \rangle \right)^2 \right] \\
&= \mathbb{E}\left[\sum_{i=1}^{M} \left((1 - a(i))^2 \langle U, G_i \rangle^2 + a(i)^2 \langle N, G_i \rangle^2 - 2a(i)(1 - a(i)) \langle U, G_i \rangle \langle N, G_i \rangle \right) \right]
\end{aligned}
$$

For each i, $\mathbb{E}\left[\langle N, G_i \rangle \right] = \langle \mathbb{E}[N], G_i \rangle = \langle 0, G_i \rangle = 0$, and
$\mathbb{E}\left[\langle N, G_i \rangle^2 \right] = \mathbb{E}\left[G_i^T N N^T G_i \right] = G_i^T \underbrace{\mathbb{E}\left[N N^T \right]}_{=\sigma^2 I_M} G_i = \sigma^2 G_i^T G_i = \sigma^2$, so:

$$
\mathbb{E}\left[\|U - \mathbf{D}\tilde{U}\|^2 \right] = \sum_{i=1}^{M} \left((1 - a(i))^2 \langle U, G_i \rangle^2 + \sigma^2 a(i)^2 \right) \tag{1}
$$

Minimizing this quantity is equivalent to minimizing $(1 - a(i))^2 \langle U, G_i \rangle^2 + \sigma^2 a(i)^2$ for each i. And differentiating with respect to $a(i)$, we find:
$\frac{d}{da(i)} \left((1 - a(i))^2 \langle U, G_i \rangle^2 + \sigma^2 a(i)^2 \right) = 2(a(i) - 1) \langle U, G_i \rangle^2 + 2\sigma^2 a(i)$, which is 0 for
$\boxed{a(i) = \frac{\langle U, G_i \rangle^2}{\langle U, G_i \rangle^2 + \sigma^2}}$.

Replacing $a(i)$ in (1) then yields

$$
\begin{aligned}
\mathbb{E}\left[\|U - \mathbf{D}_{inf}\tilde{U}\|^2 \right] &= \sum_{i=1}^{M} \left(\frac{\sigma^4}{(\langle U, G_i \rangle^2 + \sigma^2)^2} \langle U, G_i \rangle^2 + \sigma^2 \frac{\langle U, G_i \rangle^4}{(\langle U, G_i \rangle^2 + \sigma^2)^2} \right) \\
&= \sum_{i=1}^{M} \sigma^2 \langle U, G_i \rangle^2 \left(\frac{\sigma^2}{(\langle U, G_i \rangle^2 + \sigma^2)^2} + \frac{\langle U, G_i \rangle^2}{(\langle U, G_i \rangle^2 + \sigma^2)^2} \right) \\
&= \sum_{i=1}^{M} \frac{\sigma^2 \langle U, G_i \rangle^2}{\langle U, G_i \rangle^2 + \sigma^2}
\end{aligned}
$$

We have shown that $\boxed{\mathbb{E}\left[\|U - \mathbf{D}_{inf}\tilde{U}\|^2 \right] = \sum_{i=1}^{M} \frac{\sigma^2 \langle U, G_i \rangle^2}{\langle U, G_i \rangle^2 + \sigma^2}}$

Exercise 4.6

From the previous exercise (see (1)), we have

$$\mathbb{E}\left[\|U - \mathbf{D}\tilde{U}\|^2\right] = \sum_{i=1}^{M} \left((1-a(i))^2\langle U, G_i\rangle^2 + \sigma^2 a(i)^2\right) \tag{2}$$

Set $a(i) = \begin{cases} 1 & \text{if } |\langle U, G_i\rangle|^2 \geq c\sigma^2 \\ 0 & \text{otherwise.} \end{cases}$ for some $c > 1$.

- If $|\langle U, G_i\rangle|^2 \geq c\sigma^2$, then $\min(|\langle U, G_i\rangle|^2, c\sigma^2) = c\sigma^2$, and $a(i) = 1$ so $((1-a(i))^2\langle U, G_i\rangle^2 + \sigma^2 a(i)^2) = \sigma^2 \overset{c\geq 1}{\leq} c\sigma^2 = \min(|\langle U, G_i\rangle|^2, c\sigma^2)$.
- If $|\langle U, G_i\rangle|^2 < c\sigma^2$, then $\min(|\langle U, G_i\rangle|^2, c\sigma^2) = |\langle U, G_i\rangle|^2$, and $a(i) = 0$, so $((1-a(i))^2\langle U, G_i\rangle^2 + \sigma^2 a(i)^2) = |\langle U, G_i\rangle|^2 = \min(|\langle U, G_i\rangle|^2, c\sigma^2)$.

We have shown that $\forall i$, $((1-a(i))^2\langle U, G_i\rangle^2 + \sigma^2 a(i)^2) \leq \min(|\langle U, G_i\rangle|^2, c\sigma^2)$, so by summing over i, and using (2), we have

$$\boxed{\mathbb{E}\left[\|U - \mathbf{D}\tilde{U}\|^2\right] \leq \sum_{i=1}^{M} \min(|\langle U, G_i\rangle|^2, c\sigma^2) \text{ for } c > 1}$$

.

Notice that the inequality comes from the upper-bound $\sigma^2 \leq c\sigma^2$ in the case $|\langle U, G_i\rangle|^2 \geq c\sigma^2$. So when $c = 1$, all inequalities become equalities, that is:

$$\boxed{\mathbb{E}\left[\|U - \mathbf{D}\tilde{U}\|^2\right] = \sum_{i=1}^{M} \min(|\langle U, G_i\rangle|^2, c\sigma^2) \text{ for } c = 1}$$

.

Exercise 4.7

DCT:

The DCT can be rewritten $Y = AX$ where A is an $N \times N$ matrix with coefficients given by $A_{k,l} = 2\alpha_k \cos\left(\pi\left(l + \frac{1}{2}\right)\frac{k}{N}\right)$ $(0 \leq k, l \leq N-1)$, where $\alpha_k = \begin{cases} \sqrt{1/(4N)} & \text{if } k = 0 \\ \sqrt{1/(2N)} & \text{otherwise} \end{cases}$.

Saying that the DCT is an isometry is equivalent to saying that A is an isometry, i.e., that $A^T A = I_N$. For $0 \leq k, l \leq N-1$, we have:

$$\begin{aligned}
(A^T A)_{k,l} &= \sum_{m=1}^{N} (A^T)_{k,m} A_{m,l} \\
&= \sum_{m=0}^{N-1} A_{m,k} A_{m,l} \\
&= 4 \sum_{m=0}^{N-1} \alpha_m^2 \cos\left(\pi\left(k + \frac{1}{2}\right)\frac{m}{N}\right) \cos\left(\pi\left(l + \frac{1}{2}\right)\frac{m}{N}\right) \\
&= 2 \sum_{m=0}^{N-1} \alpha_m^2 \left(\cos\left(\pi\left(k + l + 1\right)\frac{m}{N}\right) + \cos\left(\pi\left(k - l\right)\frac{m}{N}\right)\right) \\
&= 2 \sum_{m=0}^{N-1} \alpha_m^2 \Re\left(\exp\left(i\pi\left(k + l + 1\right)\frac{m}{N}\right)\right) + 2 \sum_{m=0}^{N-1} \alpha_m^2 \Re\left(\exp\left(i\pi\left(k - l\right)\frac{m}{N}\right)\right) \\
&= 4 \underbrace{\alpha_0^2}_{=1/(4N)} + 2 \sum_{m=1}^{N-1} \underbrace{\alpha_m^2}_{=1/(2N)} \Re\left(\exp\left(i\pi\left(k + l + 1\right)\frac{m}{N}\right)\right) + 2 \sum_{m=1}^{N-1} \underbrace{\alpha_m^2}_{=1/(2N)} \Re\left(\exp\left(i\pi\left(k - l\right)\frac{m}{N}\right)\right) \\
&= \frac{1}{N} + \frac{1}{N}\Re\left(\sum_{m=1}^{N-1} \exp\left(i\pi\left(k + l + 1\right)\frac{m}{N}\right)\right) + \frac{1}{N}\Re\left(\sum_{m=1}^{N-1} \exp\left(i\pi\left(k - l\right)\frac{m}{N}\right)\right)
\end{aligned}$$

Then, we use the formula $\sum_{m=1}^{N-1} a^m = \begin{cases} \dfrac{a - a^N}{1 - a} & \text{if } a \neq 1 \\ N & \text{otherwise} \end{cases}$.

- We have $0 \leq k, l \leq N-1$, so $\frac{1}{N} \leq \frac{k+l+1}{N} \leq 2 - \frac{1}{N}$, so the first sum corresponds to $a \neq 1$, that is:

  $$\sum_{m=1}^{N-1} \exp\left(i\pi\left(k + l + 1\right)\frac{m}{N}\right) = \frac{\exp\left(i\pi\left(k + l + 1\right)\frac{1}{N}\right) - \exp\left(i\pi\left(k + l + 1\right)\right)}{1 - \exp\left(i\pi\left(k + l + 1\right)\frac{1}{N}\right)}$$

  We then simplify by $\exp\left(i\pi\left(k + l + 1\right)\frac{1}{2N}\right)$:

  $$\begin{aligned}
  \sum_{m=1}^{N-1} \exp\left(i\pi\left(k + l + 1\right)\frac{m}{N}\right) &= \frac{\exp\left(i\pi\left(k + l + 1\right)\frac{1}{2N}\right) - \exp\left(i\pi\left(k + l + 1\right)\left(1 - \frac{1}{2N}\right)\right)}{\exp\left(-i\pi\left(k + l + 1\right)\frac{1}{2N}\right) - \exp\left(i\pi\left(k + l + 1\right)\frac{1}{2N}\right)} \\
  &= \frac{\exp\left(i\pi\left(k + l + 1\right)\frac{1}{2N}\right) - \exp\left(i\pi\left(k + l + 1\right)\left(1 - \frac{1}{2N}\right)\right)}{-2i\sin\left(\pi\left(k + l + 1\right)\frac{1}{2N}\right)}
  \end{aligned}$$

  For $k + l$ even, $k + l + 1$ is odd, and the numerator becomes $\exp\left(i\pi\left(k + l + 1\right)\frac{1}{2N}\right) + \exp\left(-i\pi\left(k + l + 1\right)\frac{1}{2N}\right) = 2\cos\left(\pi\left(k + l + 1\right)\frac{1}{2N}\right)$, so the sum is an imaginary number, with real part 0.

For $k+l$ odd, $k+l+1$ is even, and the numerator becomes $\exp\left(i\pi\left(k+l+1\right)\frac{1}{2N}\right) - \exp\left(-i\pi\left(k+l+1\right)\frac{1}{2N}\right) = 2i\sin\left(\pi\left(k+l+1\right)\frac{1}{2N}\right)$, so the sum is equal to -1.

- For the second sum, if $k=j$ then the sum is equal to $N-1$.

  If $k \neq l$, then we have a geometric sum with $a \neq 1$, and

$$\sum_{m=1}^{N-1} \exp\left(i\pi\left(k-l\right)\frac{m}{N}\right) = \frac{\exp\left(i\pi\left(k-l\right)\frac{1}{N}\right) - \exp\left(i\pi\left(k-l\right)\right)}{1-\exp\left(i\pi\left(k-l\right)\frac{1}{N}\right)}$$
$$= \frac{\exp\left(i\pi\left(k-l\right)\frac{1}{2N}\right) - \exp\left(i\pi\left(k-l\right)\left(1-\frac{1}{2N}\right)\right)}{-2i\sin\left(\pi\left(k-l\right)\frac{1}{2N}\right)}$$

  For $k-l$ even (which is equivalent to $k+l$ even), the numerator becomes $2i\sin\left(\pi\left(k-l\right)\frac{1}{2N}\right)$ and the sum is -1.

  For $k-l$ odd, the numerator becomes $2\cos\left(\pi\left(k-l\right)\frac{1}{2N}\right)$ and the real part of the sum is 0.

So when $k \neq l$, the real part of the sums are -1 and 0, and by the above calculation we have $(A^TA)_{k,l} = 0$.
For $k=l$, $k+l=2k$ is even so the real part of the first sum is 0, and the second sum is $N-1$, which gives $(A^TA)_{k,l} = 1$.

We have shown $\boxed{A^TA = I_N}$, so $\boxed{\text{DCT defines an isometry}}$.

IDCT:

The IDCT can be rewritten $X = BY$ where B is an $N \times N$ matrix with coefficients given by $B_{k,l} = 2\tilde{\beta}_l \cos\left(\pi\left(k+\frac{1}{2}\right)\frac{l}{N}\right)$ $(0 \leq k, l \leq N-1)$, where $\tilde{\beta}_l = \begin{cases} \sqrt{1/(4N)} & \text{if } l=0 \\ \sqrt{1/(2N)} & \text{otherwise} \end{cases}$.
And recall that $A_{l,k} = 2\alpha_l \cos\left(\pi\left(k+\frac{1}{2}\right)\frac{l}{N}\right)$ with $\alpha_l = \tilde{\beta}_l$, so $B_{k,l} = A_{l,k}$ and $B = A^T$, and since $A^T = A^{-1}$, $B = A^{-1}$.
So $\boxed{\text{DCT and IDCT are inverse of each other.}}$

And $B^TB = (A^{-1})^TA^{-1} = (A^T)^{-1}A^{-1} = (A^{-1})^{-1}A^{-1} = AA^{-1} = I_N$, so B is orthogonal, and $\boxed{\text{IDCT defines an isometry}}$.

Exercise 4.8

Let $f(\alpha) = \sum\limits_k \alpha_k^2 \sigma_k^2$, and $g(\alpha) = \sum\limits_k \alpha_k$.
The optimization problem reads

$$\begin{cases} \min\limits_{\alpha} \ f(\alpha) \\ \text{s.t. } \ g(\alpha) = 1, \alpha \succeq 0 \end{cases}$$

f and g are convex, so this is a convex optimization problem. Note that the set $\{\alpha \mid \alpha \succeq 0, g(\alpha) = 1\}$ is a non-empty compact set, and f is continuous, so there exists an optimal α.

The Lagrangian reads $\mathcal{L}(\alpha; \lambda) = f(\alpha) + \lambda(1 - g(\alpha))$.

Slater's conditions hold, so there exists λ optimal for the dual problem, and using the KKT condition, we have $\nabla_\alpha \mathcal{L}(\alpha; \lambda) = 0$ for optimal α, λ, which is equivalent to $\forall k, 2\sigma_k^2 \alpha_k - \lambda = 0$, or $\boxed{\forall k, 2\alpha_k \sigma_k^2 = \lambda}$.

Exercise 4.9

By Parseval theorem, for a patch X_k, we have $\mathrm{Var}(X_k) = \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^2\right] = \sigma^2 \sum_j (\rho_{P_k})_j^2$, which can be rewritten $\sigma_k^2 = \sigma^2 \|\rho_{P_k}\|^2$.

And since $\alpha_k = \dfrac{\sigma_k^{-2}}{\sum_j \sigma_j^{-2}}$, we have $\alpha_k = \dfrac{\sigma^{-2}\|\rho_{P_k}\|^{-2}}{\sigma^{-2}\sum_j \|\rho_{P_j}\|^{-2}}$, that is $\boxed{\alpha_k = \dfrac{\|\rho_{P_k}\|^{-2}}{\sum_j \|\rho_{P_j}\|^{-2}}}$.