Course 1 assignment Denys Sikorskyi

Exercise 4.1

There is a discrete variable x on N with distribution p_n.

$$Ex = \sum_n n p_n = \sum_n n \cdot \frac{\lambda^n e^{-\lambda}}{n!} = \lambda e^{-\lambda} (\sum_n \frac{\lambda^{n-1}}{(n-1)!} =$$

$$= \lambda e^{-\lambda} \underbrace{\sum_n \frac{\lambda^{n-1}}{(n-1)!}}_{= e^{\lambda}} = \lambda. \quad \sigma^2(x) = Ex^2 - (Ex)^2$$

$$Ex^2 = \sum_{n \geq 0} n^2 p_n = \sum_{n \geq 0} n^2 \frac{\lambda^n e^{-\lambda}}{n!} = \lambda e^{-\lambda} \sum_{n \geq 1} n \frac{\lambda^{n-1}}{(n-1)!} = \lambda e^{-\lambda} \times$$

$$\times \sum_{n \geq 0} (n+1) \frac{\lambda^n}{n!} = \lambda \underbrace{e^{-\lambda} \sum_{n \geq 0} n \frac{\lambda^n}{n!}}_{Ex} + \lambda = \lambda^2 + \lambda$$

$$Ex^2 - (Ex)^2 = \lambda^2 + \lambda - \lambda^2 =$$

$$= \lambda \Rightarrow Ex = \sigma^2(x) = \lambda.$$

Exercise 4.2

$$x_i \sim p_n(\lambda_i), \quad \lambda = \sum_{i=1}^n \lambda_i,$$

$y = \sum_{i=1}^n x_i \in N$. For $X \sim p(\lambda)$ $G_x(t) = E t^x = \sum_{n \geq 0} \frac{t^n \lambda^n e^{-\lambda}}{n!} =$

$$= e^{-\lambda} \sum_{n \geq 0} \frac{(t\lambda)^n}{n!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}$$

x_i are independant

$$G_y(t) = Et^y = E t^{(\sum_{i=1}^n x_i)} = E(\prod_{i=1}^n t^{x_i}) = \prod_{i=1}^n E t^{x_i} = \prod_{i=1}^n e^{\lambda_i (t-1)} =$$

$$= e^{\sum_{i=1}^n \lambda_i (t-1)} = e^{\lambda(t-1)}. \quad G_y(t) = G_x(t) \Rightarrow y \sim P(\lambda) \Rightarrow$$

$$\Rightarrow \sum_{i=1}^n x_i \sim P(\sum_{i=1}^n \lambda_i), \quad x_i \sim P(\lambda_i), \forall i \in \{1, \dots, n\}.$$

Exercise 4.3

For noisy image $\tilde{u}$

$\tilde{u} \simeq u + g(u) \cdot n$, u - noiseless ideal image

$g(u)$ - standard deviation at the image

n - noise

$g(u)$ depends on u – image, so in order to get Gaussian noise the VST transformation is applied to stabilize s.d. Therefore, we get $f(\tilde{u}) \simeq f(u) + f'(u)\,g(u)\,n$. We consider the case of a linear variance noise $\Rightarrow (g(u))^2 = u \Leftrightarrow g(u) = \sqrt{u}$

We use the most classic VST – Anscombe transformation $f(u) = b\sqrt{u} \Rightarrow f'(u) = b\,\frac{1}{2\sqrt{u}} \Rightarrow f(\tilde{u}) \simeq f(u) + b\cdot\frac{1}{2\sqrt{u}}\cdot\sqrt{u}\cdot n \Rightarrow f(\tilde{u}) \simeq f(u) + \frac{b}{2}n \Rightarrow$ the variance is stabilized.

Therefore, $f(\tilde{u}) \simeq f(u)$ after denoising.

If we want to get the original domain of the image, the inverse transformation of VST is used $f^{-1}(y) \to \left(\frac{y}{b}\right)^2$

$f^{-1}(f(\tilde{u})) \simeq f^{-1}(f(u))$ denoised image u^* ideal image

In the end, we have $u^* \simeq u$, where u^* – obtained image after the three-step algorithm with the standard variance stabilizing transformation.

Exercise 9.5

$D_{int} = \underset{D}{\arg\min}\, E\{\|U - D_{int}\tilde{U}\|^2\}$. Note: $u_i = |\langle U, G_i\rangle|$.

D_{int} is given by $a(i) = \frac{u_i^2}{u_i^2 + \sigma^2}$

$$E(\|U - D\tilde{U}\|^2) = E\left(\sum_{i=1}^{M}(u_i - a_i\tilde{u}_i)^2\right) =$$

$$= E\left(\sum_{i=1}^{M} u_i^2 + a_i^2\tilde{u}_i^2 - 2a_i u_i \tilde{u}_i\right) = E\left(\sum_{i=1}^{M} u_i^2 + a_i^2(u_i^2 + n(i)^2 + 2a_i u_i(u_i + n_i)\right) = E\left(\sum_{i=1}^{M} \mu_i^2 + a_i^2(\mu_i^2 + n(i)^2 + 2u_i n(i)) - 2a_i u_i^2 - 2a_i \mu_i n(i)\right) = \sum_{i=1}^{M} E\left(\mu_i^2(1-a_i)^2 + a_i^2 n(i)^2 + 2u_i n(i)(a_i^2 - a_i)\right) =$$

$$= \sum_{i=1}^{M} E\left(\mu_i^2(1-a_i)^2 + a_i^2 n(i)^2 - 2u_i n(i)\, a_i(1-a_i)\right) =$$

$$= \sum_{i=1}^{M} (1-a_i)^2 \mu_i^2 + a_i^2 \underbrace{E\{n(i)^2\}}_{=\sigma^2} - 2a_i(1-a_i)\mu_i \underbrace{E\{n(i)\}}_{=0} =$$

$$= \sum_{i=1}^{M} (1-a_i)^2 \mu_i^2 + a_i^2 \sigma^2$$

The MSE given by a sum of positive terms, so it can be minimized with regardtoeach term. $f: x \to (1-x)^2 a^2 + x^2 b^2 \,|\, a=\mu_i; b=\sigma$

$f'(x) = -2(1-x)a^2 + 2xb = 0 \iff x b^2 = a^2 - xa^2 \Rightarrow x(b^2+a^2) = a^2 \iff$

$\iff x = \frac{a^2}{a^2+b^2} \Rightarrow a_i = \frac{\mu_i^2}{\mu_i^2+\sigma^2} = \frac{|\langle u, G_i\rangle|^2}{|\langle u, G_i\rangle|^2+\sigma^2}$

$E\{\|u - D_{int}\tilde{u}\|^2\} = \sum_{i=1}^{M} \frac{|\langle u, G_i\rangle|^2 \sigma^2}{|\langle u, G_i\rangle|^2+\sigma^2}$ — ?

$E(\|u - D_{int}\tilde{u}\|^2) = \sum_{i=1}^{M} (1-a_i)\mu_i^2 + a_i^2\sigma^2 = \sum_{i=1}^{M} \left(1 - \frac{\mu_i^2}{\mu_i^2+\sigma^2}\right)^2 \mu_i^2 +$

$+ \left(\frac{\mu_i^2}{\mu_i^2+\sigma^2}\right)^2 \sigma^2 = \sum_{i=1}^{M} \frac{\sigma^4 \mu_i^2}{(\mu_i^2+\sigma^2)^2} + \frac{\mu_i^4 \sigma^2}{(\mu_i^2+\sigma^2)^2} = \sum_{i=1}^{M} \frac{\mu_i^2\sigma^2(\mu_i^2+\sigma^2)}{(\mu_i^2+\sigma^2)^2} =$

$= \sum_{i=1}^{M} \frac{\mu_i^2\sigma^2}{\mu_i^2+\sigma^2} \Rightarrow E(\|u - D_{int}\tilde{u}\|^2) = \sum_{i=1}^{M} \frac{|\langle u, G_i\rangle|^2\sigma^2}{|\langle u, G_i\rangle|^2+\sigma^2}$

Exercise 9.6

Projection operator $a_i \in \{0,1\}$. The projection operator that minimizes the MSE under the constraint $a_i \in \{0,1\}$ $\forall i$, is given by:

$$a(i) = \begin{cases} 1 & \text{if } |\langle u, G_i\rangle|^2 \geq c\cdot\sigma^2, \quad c \geq 1 \\ 0 & \text{, otherwise} \end{cases}$$

We show that corresponding MSE satisfies:

$E(\|u - D_{int}\tilde{u}\|^2) \leq \sum \min(\langle u, G_i\rangle^2, c\sigma^2)$

and that inequality becomes an equality for $c=1$.

Previously in 4.5 was proved that $E(\|u - D_{int}\tilde{u}\|^2) =$

$= \sum_{i=1}^{M} ((1-a_i)^2 u_i^2 + a_i^2 \sigma^2)$

$c > 1 \Rightarrow$ if $|\langle u, G_i \rangle| \geq c\sigma^2 \Rightarrow a_i = 1$.

$$\Rightarrow \begin{cases} E(\|u - D\tilde{u}\|^2) = \sum_{i=1}^{M} \sigma^2 < \sum_{i=1}^{n} c\sigma^2 \\ \min\{|u_i|^2, c\sigma^2\} = c\sigma^2 \end{cases} \Rightarrow E(\|u - D\tilde{u}\|^2) <$$

$< \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\}$.

if $|\langle u, G_i \rangle|^2 < c\sigma^2 \Rightarrow a_i = 0$

$$\Rightarrow \begin{cases} E(\|u - D\tilde{u}\|^2) = \sum_i u_i^2 = \sum_i |\langle u, G_i \rangle|^2 \\ \min\{|\langle u, G_i \rangle|^2, c\sigma^2\} = |\langle u, G_i \rangle|^2 \end{cases} \Rightarrow E(\|u - D\tilde{u}\|^2) =$$

$= \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\} \Rightarrow$ when $c > 1$, then

$E(\|u - D\tilde{u}\|^2) \leq \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\}$

$c = 1 \Rightarrow$ if $|\langle u, G_i \rangle|^2 \geq \sigma^2 \Rightarrow a_i = 1$

$$\Rightarrow \begin{cases} E(\|u - D\tilde{u}\|) = \sum_i \sigma^2 \\ \min\{|\langle u, G_i \rangle|^2, c\sigma^2\} = c\sigma^2 = \sigma^2 \end{cases} \Rightarrow E(\|u - D\tilde{u}\|^2) =$$

$= \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\}$

if $|\langle u, G_i \rangle|^2 < \sigma^2 \Rightarrow a_i = 0$

$$\Rightarrow \begin{cases} E(\|u - D\tilde{u}\|^2) = \sum_i |\langle u, G_i \rangle|^2 \\ \min\{|\langle u, G_i \rangle|^2, c\sigma^2\} = |\langle u, G_i \rangle|^2 \end{cases}$$

$\Rightarrow E(\|u - D\tilde{u}\|^2) = \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\} \Rightarrow$

if $c = 1 \Rightarrow E(\|u - D\tilde{u}\|^2) = \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\}$.

We proved that $E(\|u - D_{int}\tilde{u}\|^2) \leq \sum_i \min\{|\langle u, G_i \rangle|^2, c\sigma^2\}$,

and we have equality if $c = 1$.

Exercise 4.2

Definition of DCT:

$\forall k \in \{0,\dots,N-1\},\ j \in \{0,\dots,N-1\}$, where $\alpha_k = \begin{cases} \sqrt{\frac{1}{4N}}, & k=0 \\ \sqrt{\frac{1}{2N}}, & k=1,\dots,N-1 \end{cases}$

$y_k = \alpha_k \cdot 2 \cdot \sum_{j=0}^{N-1} x_j \cos\left(\pi\left(j+\frac{1}{2}\right)\frac{k}{N}\right)$,

We show that the DCT is an isometry:

DCT: $y \to Ax$, where $A = (A_{kj})_{0 \le k,j \le N-1}$

$A_{kj} = 2\alpha_k \cos\left(\pi\left(j+\frac{1}{2}\right)\frac{k}{N}\right)$

The DCT is a linear transformation. So, it is an isometry iff A is orthogonal.

We show that $A^T A = I$.

$$(A^TA)_{ij} = \sum_{k=0}^{N-1} (A^T)_{ik} A_{kj} = \sum_{k=0}^{N-1} A_{ki} A_{kj} = \sum_{k=0}^{N-1} \left[2\alpha_k \cos\left(\pi\left(i+\frac{1}{2}\right)\frac{k}{N}\right)\right]$$

$$= 4\sum_{k=0}^{N-1} \frac{1}{2}\alpha_k^2 \left[\cos\left(\frac{\pi k}{N}(i+j+1)\right) + \cos\left(\frac{\pi k}{N}(i-j)\right)\right] = 2\sum_{k=0}^{N-1} \alpha_k^2 \left[Re\left(e^{\frac{i\pi k}{N}(i+j+1)} + e^{\frac{i\pi k}{N}(i-j)}\right)\right] = 2Re\left(\frac{2}{4N} + \frac{1}{2N}\sum_{k=0}^{N-1}\left(e^{\frac{i\pi k}{N}(i+j+1)} + e^{\frac{i\pi k}{N}(i-j)}\right)\right) =$$

$$= Re\left(\frac{1}{N} + \frac{1}{N}\sum_{k=1}^{N-1}\left(e^{\frac{i\pi k(i+j+1)}{N}} + e^{\frac{i\pi k}{N}(i-j)}\right)\right) = \frac{1}{N} Re\left(1 + \sum_{k=1}^{N-1}\left(e^{\frac{i\pi k(i+j+1)}{N}} + e^{\frac{i\pi k}{N}(i-j)}\right)\right)$$

if $i \neq j$

$$(A^TA)_{ij} = \frac{1}{N} Re\left(1 + \left(\frac{e^{\frac{i\pi(i+j+1)}{N}}\left(1 - e^{\frac{i\pi(i+j+1)(N-1)}{N}}\right)}{1 - e^{\frac{i\pi(i+j+1)}{N}}}\right) + \left(\frac{e^{\frac{i\pi(i-j)}{N}}\left(1 - e^{i\pi(i-j)\left(\frac{N-1}{N}\right)}\right)}{1 - e^{\frac{i\pi(i-j)}{N}}}\right)\right) = \frac{1}{N} Re\left(1 + \left(\frac{e^{\frac{i\pi(i+j+1)}{N}} - e^{i\pi(i+j+1)}}{1 - e^{\frac{i\pi(i+j+1)}{N}}} + \frac{e^{\frac{i\pi(i-j)}{N}} - e^{i\pi(i-j)}}{1 - e^{\frac{i\pi(i-j)}{N}}}\right)\right)$$

if $i+j+1 \vdots 2 \Rightarrow i+j \not\vdots 2 \Rightarrow i-j \not\vdots 2 \Rightarrow \begin{cases} e^{i\pi(i+j+1)} = 1 \\ e^{i\pi(i-j)} = -1 \end{cases} \Rightarrow$

$$\Rightarrow (A^TA)_{ij} = \frac{1}{N} Re\left(1 - 1 + \frac{1+e^{\frac{i\pi(i-j)}{N}}}{1-e^{\frac{i\pi(i-j)}{N}}}\right) = \frac{1}{N} Re\left(\underbrace{\frac{2\cos(\frac{\pi(i-j)}{2N})}{2i\sin(\frac{\pi(i-j)}{2N})}}_{\text{imaginary}}\right) = 0$$

if $i+j+1 \not\vdots 2 \Rightarrow i+j \vdots 2 \Rightarrow i-j \vdots 2 \Rightarrow \begin{cases} e^{i\pi(i+j+1)} = -1 \\ e^{i\pi(i-j)} = 1 \end{cases} \Rightarrow$

$$\Rightarrow (A^TA)_{ij} = \frac{1}{N} Re\left(1 + \frac{1+e^{\frac{i\pi(j+i+1)}{N}}}{1-e^{\frac{i\pi(i+j+1)}{N}}} - 1\right) = \frac{1}{N} Re\left(\underbrace{\frac{2\cos(\frac{\pi(i+j+1)}{2N})}{2i\sin(\frac{\pi(i+j+1)}{2N})}}_{\in Im}\right) = 0$$

Result, $\forall i \neq j \ (A^TA)_{ij} = 0$

if $i = j$

$$(A^TA)_{jj} = 2Re\left(\sum_{k=0}^{N} \alpha_k^2 \left(e^{\frac{i\pi k}{N}(2j+1)} + 1\right)\right) = 2Re\left(\frac{1}{4N} \cdot 2 + \sum_{k=1}^{N-1} \left(\frac{1}{2N} e^{\frac{i\pi k}{N}(2j+1)} + \frac{N-1}{2N}\right)\right) = 2Re\left(\frac{1}{2} + \frac{1}{2N} - \frac{1}{2N} + \sum_{k=1}^{N-1} e^{\frac{i\pi k}{N}(2j+1)}\right) =$$

$$= 2Re\left(\frac{1}{2} + \frac{1}{2N} e^{\frac{i\pi}{N}(2j+1)} \left(\frac{1-e^{\frac{i\pi}{N}(N-1)(2j+1)}}{1-e^{\frac{i\pi}{N}(2j+1)}}\right)\right) =$$

$$= 1 + \frac{1}{N} Re\left(\frac{e^{\frac{i\pi}{N}(2j+1)} - \underbrace{e^{i\pi(2j+1)}}_{= -1 \text{ cause } 2j+1 \not\vdots 2}}{1-e^{\frac{i\pi}{N}(2j+1)}}\right) =$$

$$= 1 + \frac{1}{N} Re\left(\frac{1+e^{\frac{i\pi}{N}(2j+1)}}{1-e^{\frac{i\pi}{N}(2j+1)}}\right) = 1 + \frac{1}{N} Re\left(\underbrace{\frac{2\cos(\frac{\pi}{2N}(2j+1))}{2i\sin(\frac{\pi}{2N}(2j+1))}}_{\in Im}\right) =$$

$$= 1 \Rightarrow (A^TA)_{jj} = 1, \ \forall j.$$

Thus, $A^TA = I$, so DCT is an isometry.

Using same calculations we prove that IDCT is an isometry $B^TB = I$, $\left(B_{jk} = 2\beta_k \cos\left(\pi\left(j+\frac{1}{2}\right)\frac{k}{N}\right)\right)$.

We have

$$x_j = \beta_0 y_0 + \sum_{k=1}^{N-1} \beta_k 2 y_k \cos\left(\pi\left(j+\frac{1}{2}\right)\frac{k}{N}\right) \text{ with } \beta_k =$$

$$= \begin{cases} \sqrt{\frac{1}{N}}, & k=0 \\ \sqrt{\frac{1}{2N}}, & k=1, \dots N-1 \end{cases}$$

$x_j = \sum_{k=0}^{N-1} \beta_k' 2 y_k \cos(\pi(j+\frac{1}{2})\frac{k}{N})$, but $\beta_0' = \sqrt{\frac{1}{4N}}$ and $\forall k \in \{1, \dots, N-1\}$ $\beta_k' = \sqrt{\frac{1}{2N}}$

At last, let's show that DCT and IDCT are inverse of each other. We computed that $A^T A = I = \cancel{AA^T} \Rightarrow$
$\Rightarrow A^{-1} = A^T$. Let's show that $A^T = B$

$$\forall k, \forall j : \begin{cases} A_{kj} = 2\alpha_k \cos(\pi(j+\frac{1}{2})\frac{k}{N}) \\ B_{jk} = 2\beta_k' \cos(\pi(j+\frac{1}{2})\frac{k}{N}) \end{cases}$$

So, $\forall k$, $\alpha_k = \beta_k' \Rightarrow \forall j,k\ A_{kj} = B_{jk} \Rightarrow B_{jk} = A^T_{jk} \Rightarrow$
$\Rightarrow B = A^T = A^{-1}$.

Exercise 4.8

We show that the constrained optimization problem:

$$\underset{\sum \alpha_k = 1}{\operatorname{argmin}} \sum_k \alpha_k^2 \underbrace{E[(p_k - E[p_k])^2]}_{\sigma_k^2} \quad (1)$$

implies the existance of some $\lambda \in \mathbb{R}$ such that:

$$2\alpha_k \sigma_k^2 = \lambda, \forall k$$

We can write the problem (1) as:

$$\underset{\alpha}{\operatorname{argmin}} \sum_{k=1}^{K^2} \alpha_k^2 \sigma_k^2$$

subject to $\begin{cases} g(\alpha) = \sum_{k=1} \alpha_k - 1 \\ \forall k \in \{1, \dots, K^2\}, \alpha_k \geq 0 \end{cases}$

We denote by $\begin{cases} f(\alpha) = \sum_{k=1}^{K} \alpha_k^2 \sigma^2 \\ C = \{\alpha \mid \sum \alpha_k = 1, \alpha_k \geq 0\ \forall k\} \end{cases}$

$\begin{cases} f \text{ is convex and l.s.c on } C \\ C \text{ is convex and compact} \end{cases} \Rightarrow$

$\Rightarrow \exists$ attainable minimizer: $\exists \hat{\alpha} \in C$ such that $f(\hat{\alpha}) = \inf f(\alpha)$

From Slater's equality constraint, the problem is convex and feasible: $\begin{cases} C \neq \emptyset \\ \forall k \in \{1,\dots,k^2\},\ \alpha_k = \frac{1}{k^2} \in C. \end{cases}$

We consider the lagrangian function:

$$\mathcal{L}(\alpha,\lambda) = \sum_k \alpha_k^2 \sigma^2 + \left(1 - \sum_k \alpha_k\right)\lambda$$

Using the KKT conditions:

$(\hat{\alpha},\lambda)$ is a saddle point $\iff \nabla_\alpha \mathcal{L}(\hat{\alpha},\hat{\lambda}) = 0 \iff \forall k,\ \frac{\partial \mathcal{L}(\hat{\alpha},\hat{\lambda})}{\partial \alpha_k} = 0$

$\iff 2\sigma_k^2 \hat{\alpha}_k = \hat{\lambda} \Rightarrow \forall k,\ 2\hat{\alpha}_k \sigma_k^2 = \lambda.$

Therefore, $\exists \lambda \in \mathbb{R}$ such that $2\alpha_k \sigma_k^2 = \lambda$.

Exercise 4.9

Let's evaluate the variance of each patch element. det N_{P_k} — number of $\neq 0$ coefs DCT in the k-th patch after hard thresholding. From Parseval's formula for a patch k $var(X_k) = \sigma_k^2 = \sigma^2 \sum_{j=1}^{N} a_j^2$, a_j – coefs of thresholding.

Hard thresholding $\Rightarrow a_j \in \{0,1\}\ \forall j \in \{1,\dots,N\} \Rightarrow \sigma_k^2 = \sigma^2 \underbrace{\sum_{j=1}^{N} a_j}_{N_{P_k}} = \sigma^2 N_{P_k}$

The optimal weights are: $\alpha_k = \frac{\sigma_k^{-2}}{\sum_k \sigma_k^{-2}},\ \forall k \Rightarrow \forall k\ \alpha_k = \frac{N_{P_k}^{-1}}{\sum_j N_{P_j}^{-1}}$

Wiener filtering: the coefficients are given by Wiener coefficients $\Rightarrow \forall k,\ a_k = e_{P_k} \Rightarrow \sigma_k^2 = \sigma^2 \sum_{j=1}^{N} a_j = \sigma^2 \underbrace{\sum_j e_{P_j}^2}_{=\|e_{P_k}\|^2} =$

$= \sigma^2 \|e_{P_k}\|^2$ Or, the optimal weights are given by $\alpha_k = \frac{\sigma_k^{-2}}{\sum_k \sigma_k^{-2}},\ \forall k \Rightarrow$

$\Rightarrow \forall k,\ \alpha_k = \frac{\sigma^{-2}\|e_{P_k}\|^2}{\sum_j \sigma^{-2}\|e_{P_j}\|^{-2}} \Rightarrow \forall k,\ \alpha_k = \frac{\|e_{P_k}\|^{-2}}{\sum_j \|e_{P_j}\|^{-2}}$. So, the formula of aggregation weight in both cases of hard thresholding and Wiener function were demonstrated!