Exercise 3.1 Optimal convex combination of a blind-spot network and the noisy image.

Suppose that we train a blind-spot network $\mathcal{F}$. Images are contaminated with noise that (i) is unbiased, i.e. $E_v\{v|u\} = u$ for all u; (ii) is spatially independent i.e. v_x and v_y, for $x \neq y$, are conditionally independent given u.

1. Show that the MSE of $\mathcal{F}^\lambda(v)$ can be decomposed as $E_{v,u}\{\|\mathcal{F}^\lambda(v)-u\|^2\} = \lambda^2 E_{v,u}\{\|\mathcal{F}^\lambda(v)-u\|^2\} + (1-\lambda)^2 \cdot E_{v,u}\{\|v-u\|^2\} + 2\lambda(1-\lambda) E_{v,u}\{\langle \mathcal{F}(v)-u, v-u \rangle\}$.

$E_{v,u}\{\|\mathcal{F}^\lambda(v)-u\|^2\} = E_{v,u}\{\|\lambda F(v) + (1-\lambda)v - u\|^2\} =$
$= E_{v,u}\{\|\lambda(F(v)-u) + (1-\lambda)(v-u)\|^2\} = E_{v,u}\{\lambda^2\|F(v)-u\|^2 +$
$+ (1-\lambda)^2\|v-u\|^2 + 2\lambda(1-\lambda)\langle F(v)-u, u-v\rangle\}$ ⊜ we can split by addition sign ⊜ $\lambda^2 E_{v,u}\{\|\mathcal{F}^\lambda(v)-u\|^2\} + (1-\lambda)^2 \cdot E_{v,u}\{\|v-u\|^2\} + 2\lambda(1-\lambda) E_{v,u}\{\langle \mathcal{F}(v)-u, v-u \rangle$ – the expression that we wanted to achieve.

2. Show that $E_{v,u}\{\langle \mathcal{F}(v)-u, v-u\rangle\} = 0$

$E_{v,u}\{\langle \mathcal{F}(v)-u, v-u\rangle\} = E_u\{E_v\{\langle \mathcal{F}(v)-u, u-v\rangle | u\}\} =$
$= E_u\{E_v\{\sum_{j\in J} \langle F(v)_j - u_j, u_j - v_j\rangle | u\}\} = E_u \sum_{j \in J} E_v\{\langle F(v)_j -$
$- u_j, u_j - v_j\rangle | u\}$ $^{(*)}$ Let's prove that inner expectation is zero

$E_v\{\langle F(v)_j - u_j, u_j - v_j = E_{v_j^c}\{E_{v_j}\{\langle F(v)_j - u_j, u_j - v_j | v_j^c, u\rangle | u\}$
$= E_{v_j^c}\{\langle F(v)_j - u_j, E_{v_j}\{u_j - v_j | v_j^c, u\}\rangle | u\}$ ⊜ $F(v)_j$ is unchanged with v_j, cause F is J-invariant

$(=)\ E_{v_j}\{\langle F(v)_j - x_j, x_j - E\{v_j | x\}\rangle | x\} \ (=)\ 0$

Images are noised with spatially independent noise and unbiased noise, so $E\{v_j | x\} = x_j$ $(=)$ 0.

We have that each part of summation $(*)$ is 0, then

$$E_{v,x}\{\langle F(v) - x, x - v\rangle\} = 0$$

3. Deduce that λ^* that minimizes the MSE is given by

$$\lambda^* = \frac{E_x\{V\{v|x\}\}}{E_{v,x}\{\|F(v) - x\|^2\} + E_x\{V\{v|x\}\}}$$

Using result from part 2 we can exclude last expression of MSE and get

$$E_{v,x}\{\|F^\lambda(v) - x\|^2\} = \lambda^2 E_{v,x}\{\|F(v) - x\|^2\} + (1-\lambda)^2 E_{v,x}\{\|v - x\|^2\} \to \min_\lambda \Rightarrow$$

$$\Rightarrow \frac{\partial E_{v,x}\{\|F^\lambda(v) - x\|^2\}}{\partial \lambda} = 2\lambda E_{v,x}\{\|F(v) - x\|^2\} - 2(1-\lambda) E_{v,x}\{\|v - x\|^2\} = 0 \Rightarrow$$

$$\Rightarrow \lambda = \frac{E_{v,x}\{\|v - x\|^2\}}{E_{v,x}\{\|F(v) - x\|^2\} + E_{v,x}\{\|v - x\|^2\}} \ (=)$$

$(=)$ $E_{v,x}\{\|v - x\|^2\} = E_x\{V\{v | x\}\} \Rightarrow$ We have desired expression:

$$\lambda^* = \frac{E_x\{V\{v|x\}\}}{E_{v,x}\{\|F(v) - x\|^2\} + E_x\{V\{v|x\}\}}$$

4. Suppose now that the noise has a variance $V\{v|u\} = d\sigma^2$, for all $u \in \mathbb{R}^d$. Use Proposition 3.4 and express λ^* in terms of σ^2 and the self-supervised risk $R_{N2S}(\mathcal{F})$.

We have from Proposition 3.5: $E_v\{\|F(v) - v\|^2\} =$
$= E_{v,u}\{\|\mathcal{F}(v) - u\|^2\} + E_u\{V\{v|u\}\}$. Also,
$R_{N2S}(\mathcal{F}) = E_v\{\|F(v) - v\|^2\}$. Using that

$$\lambda^* = \frac{E_u\{V\{v|u\}\}}{E_{v,u}\{\|\mathcal{F}(v) - u\|^2\} + E_u\{V\{v|u\}\}} =$$

$$= \frac{d\sigma^2}{R_{N2S}(\mathcal{F})}$$

Exercise 3.2 Bias-variance decomposition.

Given an estimator $\hat{u}(v)$ of u, where v is noisy version of u. Show that, for a given u the MSE can be expressed as follows:

$$E_v\{\|\hat{u}(v) - u\|^2\} = \overbrace{\|E_v\{\hat{u}(v)|u\} - u\|^2}^{\text{bias given } u} +$$

$$+ \underbrace{E_v\{\|\hat{u}(v) - E_v\{\hat{u}(v)|u\}\|^2|u\}}_{\text{variance given } u}$$

$E_v\{\|\hat{u}(v) - u\|^2|u\} = E_v\{\|\hat{u}(v) - E_v\{\hat{u}(v)|u\} +$
$+ E_v\{\hat{u}(v) - u\|^2\} = E_v\{\|\hat{u}(v) - E_v\{\hat{u}(v)|u\}\|^2\} +$
$+ E_v\{\|E_v\{\hat{u}(v)|u\} - u\|^2\} - 2E_v\{\langle \hat{u}(v) - E_v\{\hat{u}(v)$
$|u\}, E_v\{\hat{u}(v), u\} - u\rangle |u\}$

$= \|E_v\{\hat{u}(v)\} - u\|$

Because in the scalar product we have
$E_v\{E_v\{\hat{u}(v)|u\} - u|u\}$ – inner expectation independent of v $\Rightarrow$ the whole scalar product is equal to 0. Thus, we can finally ~~ag~~ get our expression:

$$E_v\{\|\hat{u}(v) - u\|^2|u\} = \|E_v\{\hat{u}(v)|u\} - u\|^2 + E_v\{\|\hat{u}(v) - E_v\{\hat{u}(v)|u\}\|^2|u\}.$$