General definitions

1.
$$\sigma = 0.1$$

2.
$$h: \mathbb{R}^2 o \mathbb{R}: x \mapsto h(x) = \exp\left(rac{-\|x\|^2}{\sigma^2}
ight)$$

3.
$$arphi:\mathbb{R}^2 o\mathbb{R}^{n^2}:x\mapstoarphi(x)=egin{bmatrix}h(p_1-x)\ dots\h(p_{n^2}-x)\end{bmatrix}$$

4.
$$\Phi: \mathbb{R}^{2 imes K} o \mathbb{R}^{n^2}: X \mapsto \Phi(X) = \sum_{k=1}^K arphi(x_k)$$

5.
$$f: \mathbb{R}^{2 imes K} o \mathbb{R}: X \mapsto f(X) = rac{1}{2} \left\| \Phi(X) - y
ight\|^2$$

- 6. K is the number of stars
- 7. n^2 is the number of pixels we sample (Distributed in a $n \times n$ grid)

Exercise 3

Obtain an expression for $D\varphi(x)[u]$

If we define $h_i(x) = h(p_i - x)$ for $i \in \{1, \dots, n^2\}$, then we have

$$egin{aligned} Darphi(x)[u] &= \lim_{t o 0} rac{arphi(x+tu) - arphi(x)}{t} \ &= \lim_{t o 0} rac{1}{t} egin{bmatrix} h(p_1-x-tu) - h(p_1-x) \ dots \ h(p_{n^2}-x-tu) - h(p_{n^2}-x) \end{bmatrix} = \ &= egin{bmatrix} \lim_{t o 0} rac{h(p_1-x-tu) - h(p_1-x)}{t} \ dots \ \lim_{t o 0} rac{h(p_1-x-tu) - h(p_{n^2}-x)}{t} \end{bmatrix} = \ &= egin{bmatrix} Dh(p_1-x)[-u] \ dots \ Dh(p_{n^2}-x)[-u] \end{bmatrix} = \ &= egin{bmatrix} Dh(p_1-x)[u] \ dots \ Dh(p_{n^2}-x)[u] \end{bmatrix} = egin{bmatrix} Dh_1(x)[u] \ dots \ Dh_{n^2}(x)[u] \end{bmatrix} \end{aligned}$$

It's easy to see that $\nabla h(x)=rac{-2x}{\sigma^2}\exp\left(-rac{\|x\|^2}{\sigma^2}
ight)$, and substituting $D_{h_i}(x)[u]$ for $\langle \nabla h_i(x),u \rangle$, in our previous expression for $D\varphi(x)[u]$, we observe that thanks to h_i having a gradient, we can express it as

$$Darphi(x)[u] = egin{bmatrix} Dh_1(x)[u] \ dots \ Dh_{n^2}(x)[u] \end{bmatrix} = egin{bmatrix} \langle
abla h_1(x), u
angle \ dots \ \langle
abla h_{n^2}(x), u
angle \end{bmatrix} = A(x)u$$

where

$$A: \mathbb{R}^2 o \mathbb{R}^{n^2 imes 2}: x \mapsto egin{bmatrix} \cdots
abla h_1(x) \cdots \ dots \ \ dots \ \ \ \$$

Exercise 4

Since the adjoint of a matrix A(x) is its transpose, we have that

$$A(x)^* := egin{bmatrix} dots & \cdots & dots \
abla h_1(x) & \cdots &
abla h_{n^2}(x) \ dots & \cdots & dots \ dots & \cdots & dots \ \end{pmatrix} = A(x)^T$$

Exercise 5

Recall how we defined the function f as a function of other functions

•
$$f(X) = \frac{1}{2} \|\Phi(X) - y\|^2$$

•
$$\Phi(X) = \sum_{k=1}^K \varphi(x_k)$$

$$ullet arphi(x) = egin{bmatrix} h(p_1-x) \ dots \ h(p_{n^2}-x) \end{bmatrix} = \left[h(p_i-x)
ight]_{i=1}^{n^2}$$

And finally, the definition of the dot product $\langle A,B
angle$ for any matrices $A\in\mathbb{R}^{n imes m}$ and $B\in\mathbb{R}^{n imes k}$ (n,m,k>0)

$$\langle A, B \rangle = tr(A^T B)$$

First, we recall that $D_{arphi}(x)[u] = A(x) \ u$, where $A(x) \in \mathbb{R}^{n^2 imes 2}$.

With this information, we get the following expression for

$$egin{aligned} D\Phi(X)[U] &= \sum_{k=1}^K Darphi(x_k)[u_k] = \ &= \sum_{k=1}^K A(x_k)u_k \end{aligned}$$

And thanks to this, we can find a final expression for $D_f(X)(U)$

$$egin{aligned} Df(X)[U] &= \langle \Phi(X) - y, D\Phi(X)[u_k]
angle = \ &= \left\langle \Phi(X) - y, \sum_{k=1}^K A(x_k) u_k
ight
angle = \ &= \sum_{k=1}^K \langle \Phi(X) - y, A(x_k) u_k
angle = \ &= \sum_{k=1}^K \langle A(x_k)^* (\Phi(X) - y), u_k
angle = \ &= tr \left(\left(\left[A(x_k)^* (\Phi(X) - y)
ight]_{k=1}^K
ight)^T U
ight) = \ &= \left\langle \left[A(x_k)^* (\Phi(X) - y)
ight]_{k=1}^K, U
ight
angle \end{aligned}$$

And so, thanks to the definition of gradient, we can identify the gradient of f as

$$abla f(X) = \left[A(x_k)^*(\Phi(X)-y)
ight]_{k=1}^K = egin{bmatrix} A(x_1)^*(\Phi(X)-y)\ dots\ A(x_K)^*(\Phi(X)-y) \end{bmatrix}$$