Variational Methods for Computer Vision: Solution Sheet 3

Exercise: November 13, 2019

Part I: Theory

1. (a) Suppose x^* is a local but not a global minimizer. Then there exists a $z \in \mathbb{R}^n$ with $f(z) < f(x^*)$. Consider the line segment

$$x_\lambda = \lambda z + (1-\lambda)x^*, \lambda \in (0,1).$$

By convexity we have:

$$f(x_\lambda) = f(\lambda z + (1-\lambda)x^*) \leq \lambda f(z) + (1-\lambda)f(x^*) < \lambda f(x^*) + (1-\lambda)f(x^*) = f(x^*).$$

$\Rightarrow$ Any neighbourhood of x^* contains a point x_λ with $f(x_\lambda) < f(x^*)$, which is a contradiction to the assumption.

(b) Assume that x^* is a stationary point but not a global minimizer. Then there is a $z \in \mathbb{R}^n$ with $f(z) < f(x^*)$, and

$$\begin{aligned}\langle \nabla f(x^*), z - x^* \rangle &= \lim_{\epsilon \to 0} \frac{1}{\varepsilon}(f(x^* + \varepsilon(z - x^*)) - f(x^*)) \\ &\leq \lim_{\epsilon \to 0} \frac{1}{\varepsilon}(\varepsilon f(z) + (1-\varepsilon)f(x^*) - f(x^*)) \\ &= f(z) - f(x^*) < 0.\end{aligned}$$

Thus $\langle \nabla f(x^*), z - x^* \rangle \neq 0 \Rightarrow \nabla f(x^*) \neq 0 \Rightarrow x^*$ is not a stationary point.

2. <u>f convex $\Rightarrow$ (epi f) convex:</u>

Take arbitrary (u,a), $(v,b) \in$ epi f. Then

$$f(\lambda u + (1-\lambda)v) \leq \lambda f(u) + (1-\lambda)f(v) \leq \lambda a + (1-\lambda)b\,.$$

Thus $(\lambda u + (1-\lambda)v, \lambda a + (1-\lambda)b) = \lambda(u,a) + (1-\lambda)(v,b) \in$ epi f.

<u>(epi f) convex $\Rightarrow$ f convex:</u>

Take arbitrary $x, y \in \mathbb{R}^n$ and let $a := f(x)$, $b := f(y)$. Then $(x,a), (y,b) \in$ epi f. Since epi f is convex:

$$(\lambda x + (1-\lambda)y, \lambda a + (1-\lambda)b) \in \text{epi } f\,, \quad \text{i.e.}$$

$$f(\lambda x + (1-\lambda)y) \leq \lambda a + (1-\lambda)b = \lambda f(x) + (1-\lambda)f(y).$$

This is exactly the definition of convexity of f.

3. (a) A direct calculation shows:

$$\begin{aligned}h(\lambda x + (1-\lambda)y) &= \alpha f(\lambda x + (1-\lambda)y) + \beta g(\lambda x + (1-\lambda)y) \\ &\leq \alpha\lambda f(x) + \alpha(1-\lambda)f(y) + \beta\lambda g(x) + \beta(1-\lambda)g(y) \\ &= \lambda(\alpha f(x) + \beta g(x)) + (1-\lambda)(\alpha f(y) + \beta g(y)) \\ &= \lambda h(x) + (1-\lambda)h(y).\end{aligned}$$

(b) Since $h = \max(f, g)$, we have for each x that $h(x) \geq f(x)$ and $h(x) \geq g(x)$. Thus,

$$\lambda h(x) + (1-\lambda)h(y) \geq \lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y) \quad \text{and}$$
$$\lambda h(x) + (1-\lambda)h(y) \geq \lambda g(x) + (1-\lambda)g(y) \geq g(\lambda x + (1-\lambda)y) \ ,$$

where the second "$\geq$" sign is due to convexity of f and g, respectively. Now, since both of these relations hold, we have that

$$\lambda h(x) + (1-\lambda)h(y) \geq \max(f(\lambda x + (1-\lambda)y), g(\lambda x + (1-\lambda)y)) = h(\lambda + (1-\lambda)y) \, .$$

This is exactly the definition of convexity of h.

Alternative: We see that

$$\begin{aligned} \text{epi } f \cap \text{epi } g &= \{(x,a) \mid f(x) \leq a\} \cap \{(x,a) \mid g(x) \leq a\} \\ &= \{(x,a) \mid \max\{f(x), g(x)\} \leq a\} = \text{epi } h \end{aligned}$$

Since the intersection of two convex sets is always convex, epi h is a convex set. This implies by Ex. 2 that h is also a convex function.

Now we need to proof that the intersection of two convex sets is convex (always $\lambda \in (0,1)$):

$$\begin{aligned} & S_1, S_2 \text{ convex} \\ & \Rightarrow (\forall x, y \in S_1 \colon \lambda x + (1-\lambda)y \in S_1) \wedge (\forall x, y \in S_2 \colon \lambda x + (1-\lambda)y \in S_2) \\ & \Rightarrow \big(x, y \in S_1 \wedge x, y \in S_2 \Rightarrow \lambda x + (1-\lambda)y \in S_1 \wedge \lambda x + (1-\lambda)y \in S_2\big) \\ & \Rightarrow \forall x, y \in S_1 \cap S_2 \colon \lambda x + (1-\lambda)y \in S_1 \cap S_2 \\ & \Rightarrow S_1 \cap S_2 \text{ convex.} \end{aligned}$$

(c) Counterexample: $h(x) = \min\{(x-1)^2, (x+1)^2\}$ is clearly not convex: take e.g. $x = 1$, $y = -1$ and $\lambda = \frac{1}{2}$, then

$$h(\lambda x + (1-\lambda)y) = h(0) = 1 > 0 = \lambda h(x) + (1-\lambda)h(y) \, .$$

4.

$$\begin{aligned} h''(x) &= f(g(x))'' = \big(f'(g(x))g'(x)\big)' \\ &= \underbrace{f''(g(x))}_{\geq 0} \underbrace{g'(x)g'(x)}_{\geq 0} + f'(g(x)) \underbrace{g''(x)}_{\geq 0} \end{aligned}$$

Thus $h''(x) \geq 0$ if $f'(g(x)) \geq 0$, so f being a convex non-decreasing function is a sufficient condition for the convexity of h.

Part II: Practical Exercises

1. From the lecture, we have the following condition on u:

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}u_l} = (u_l - f_l) + \lambda \sum_{\substack{l,j \\ \text{neighbours}}} (u_l - u_j) = 0$$

$$\Rightarrow \quad (1 + \lambda n_l) u_l - \lambda \sum_{\substack{l,j \\ \text{neighbours}}} u_j = f_l \, ,$$

with n_i being the number of neighbours of pixel i. Thus the Gauss-Seidel update step becomes

$$u_i^{(k+1)} = \frac{1}{1 + \lambda n_i} \left(f_i + \lambda \sum_{\substack{j \in \mathcal{N}(i) \\ j < i}} u_j^{(k+1)} + \lambda \sum_{\substack{j \in \mathcal{N}(i) \\ j > i}} u_j^{(k)} \right)$$

for the given energy.