

Variational Methods for Computer Vision: Solution Sheet 2

Exercise: November 7, 2018

Part I: Theory

1. This follows from a direct calculation:

$$\begin{aligned} ((f * k_1) * k_2)(x, y) &= \int \left(\int k_1(s) f(x - s, y - t) ds \right) k_2(t) dt \\ &= \int \int f(x - s, y - t) k_1(s) k_2(t) ds dt \\ &= \int \int f(x - s, y - t) \frac{1}{2\pi\sigma^2} \exp\left(-\frac{s^2 + t^2}{2\sigma^2}\right) ds dt \\ &= \int \int f(x - s, y - t) K(s, t) ds dt \\ &= (f * K)(x, y) \end{aligned}$$

Separability allows to implement a two-dimensional $n \times n$ filter by consecutive application of two one-dimensional filters, which can be done in $\mathcal{O}(n)$ instead of $\mathcal{O}(n^2)$.

2. (a) Since by definition the gradient of a function is the transpose of the Jacobian we have:

$$[\nabla(f \circ R)](x) = J_{f \circ R}(x)^\top = (J_f(Rx) \circ R)^\top = R^\top \circ J_f(Rx)^\top = R^\top (\nabla f)(Rx).$$

Alternative: We can also show this directly without using the chain rule in the following way. To compute the gradient, we take a look at the directional derivative of f in an arbitrary direction $R^\top h \in \mathbb{R}^2$.

$$\begin{aligned} \langle \nabla(f \circ R)(x), R^\top h \rangle &= \lim_{\varepsilon \rightarrow 0} \frac{(f \circ R)(x + \varepsilon R^\top h) - (f \circ R)(x)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{f(Rx + \varepsilon R R^\top h) - f(Rx)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{f(Rx + \varepsilon h) - f(Rx)}{\varepsilon} \\ &= \langle (\nabla f)(Rx), h \rangle \\ &= \langle (\nabla f)(Rx), R R^\top h \rangle \\ &= \langle R^\top (\nabla f)(Rx), R^\top h \rangle. \end{aligned}$$

Since the above holds for all $R^\top h$ we can conclude that $\nabla(f \circ R)(x) = R^\top (\nabla f)(Rx)$, using the following argument:

$$\begin{aligned} \langle x, h \rangle &= \langle y, h \rangle, \forall h \\ \Leftrightarrow \langle x - y, h \rangle &= 0, \forall h \\ \Rightarrow \langle x - y, x - y \rangle &= 0, \\ \Rightarrow x &= y. \end{aligned}$$

Thus we know that $\nabla(f \circ R) = R^\top \circ \nabla f \circ R$.

(b) We show the identity directly:

$$\begin{aligned}
\|\nabla(f \circ R)(x)\| &\stackrel{(a)}{=} \|R^\top(\nabla f)(Rx)\| \\
&= \sqrt{\langle R^\top(\nabla f)(Rx), R^\top(\nabla f)(Rx) \rangle} \\
&= \sqrt{\langle (\nabla f)(Rx), RR^\top(\nabla f)(Rx) \rangle} \\
&= \sqrt{\langle (\nabla f)(Rx), (\nabla f)(Rx) \rangle} \\
&= \|(\nabla f)(Rx)\|,
\end{aligned}$$

and hence $\|\nabla(f \circ R)\| = \|(\nabla f) \circ R\|$.

(c) Denoting the partial derivatives as $\partial_x f = f_x$, $\partial_y f = f_y$, $\partial_x \partial_y f = f_{yx}$, etc., (with the convention $f_x(R\mathbf{u}) := ((\partial_x f) \circ R)(x)$ and using short hand notation $\cos(\alpha) = c$, $\sin(\alpha) = s$, $\mathbf{u} = (x, y)^\top$) we compute the following:

$$\begin{aligned}
[\Delta(f \circ R)](x, y) &= \operatorname{div}(R^\top \begin{bmatrix} f_x(R\mathbf{u}) \\ f_y(R\mathbf{u}) \end{bmatrix}) \\
&= \operatorname{div} \begin{pmatrix} cf_x(R\mathbf{u}) + sf_y(R\mathbf{u}) \\ -sf_x(R\mathbf{u}) + cf_y(R\mathbf{u}) \end{pmatrix} \\
&= \partial_x(cf_x(R\mathbf{u}) + sf_y(R\mathbf{u})) + \partial_y(-sf_x(R\mathbf{u}) + cf_y(R\mathbf{u}))
\end{aligned}$$

Now we have that

$$\begin{aligned}
\nabla f_x(R\mathbf{u}) &= \begin{bmatrix} \partial_x f_x(R\mathbf{u}) \\ \partial_y f_x(R\mathbf{u}) \end{bmatrix} = R^\top \begin{bmatrix} f_{xx}(R\mathbf{u}) \\ f_{xy}(R\mathbf{u}) \end{bmatrix} = \begin{bmatrix} cf_{xx}(R\mathbf{u}) + sf_{xy}(R\mathbf{u}) \\ -sf_{xx}(R\mathbf{u}) + cf_{xy}(R\mathbf{u}) \end{bmatrix}, \\
\nabla f_y(R\mathbf{u}) &= \begin{bmatrix} \partial_x f_y(R\mathbf{u}) \\ \partial_y f_y(R\mathbf{u}) \end{bmatrix} = R^\top \begin{bmatrix} f_{yx}(R\mathbf{u}) \\ f_{yy}(R\mathbf{u}) \end{bmatrix} = \begin{bmatrix} cf_{yx}(R\mathbf{u}) + sf_{yy}(R\mathbf{u}) \\ -sf_{yx}(R\mathbf{u}) + cf_{yy}(R\mathbf{u}) \end{bmatrix},
\end{aligned}$$

and thus:

$$\begin{aligned}
&\partial_x(cf_x(R\mathbf{u}) + sf_y(R\mathbf{u})) + \partial_y(-sf_x(R\mathbf{u}) + cf_y(R\mathbf{u})) = \\
&c\partial_x f_x(R\mathbf{u}) + s\partial_x f_y(R\mathbf{u}) - s\partial_y f_x(R\mathbf{u}) + c\partial_y f_y(R\mathbf{u}) = \\
&c(cf_{xx}(R\mathbf{u}) + sf_{xy}(R\mathbf{u})) + s(cf_{yx}(R\mathbf{u}) + sf_{yy}(R\mathbf{u})) - \\
&s(-sf_{xx}(R\mathbf{u}) + cf_{xy}(R\mathbf{u})) + c(-sf_{yx}(R\mathbf{u}) + cf_{yy}(R\mathbf{u})) = \\
&c^2 f_{xx}(R\mathbf{u}) + cs f_{xy}(R\mathbf{u}) + sc f_{yx}(R\mathbf{u}) + s^2 f_{yy}(R\mathbf{u}) + \\
&s^2 f_{xx}(R\mathbf{u}) - sc f_{xy}(R\mathbf{u}) - cs f_{yx}(R\mathbf{u}) + c^2 f_{yy}(R\mathbf{u}) = \\
&\underbrace{(c^2 + s^2)}_{=1} f_{xx}(R\mathbf{u}) + \underbrace{(c^2 + s^2)}_{=1} f_{yy}(R\mathbf{u}) = ((\Delta f) \circ R)(\mathbf{u}),
\end{aligned}$$

which shows $(\Delta f) \circ R = \Delta(f \circ R)$.

3. (a) This immediately follows from the linearity of the divergence operator:

$$\begin{aligned}
\operatorname{div}(g \cdot \nabla u)(x) &= g \operatorname{div}(\nabla u)(x) \\
&= g \Delta u(x)
\end{aligned}$$

(b) We compute:

$$\begin{aligned}\operatorname{div}(g\nabla u)(x) &= \frac{\partial}{\partial x_1} \left(g \frac{\partial}{\partial x_1} u \right) (x) + \frac{\partial}{\partial x_2} \left(g \frac{\partial}{\partial x_2} u \right) (x) \\ &= \frac{\partial^2 u}{\partial x_1^2}(x)g(x) + \frac{\partial^2 u}{\partial x_2^2}(x)g(x) + \frac{\partial g}{\partial x_1}(x) \frac{\partial u}{\partial x_1}(x) + \frac{\partial g}{\partial x_2}(x) \frac{\partial u}{\partial x_2}(x) \\ &= g(x)\Delta u(x) + \langle \nabla g(x), \nabla u(x) \rangle\end{aligned}$$