Supplementary Material: Mathematical Deductions

Learning PDEs for Image Restoration via Optimal Control

Risheng Liu[1], Zhouchen Lin[2], Wei Zhang[3], and Zhixun Su[1]

[1] Dalian University of Technology, Dalian 116024, P.R. China
rsliu@mail.dlut.edu.cn, zxsu@dlut.edu.cn
[2] Microsoft Research Asia, Beijing 100190, P.R. China (corresponding author)
zhoulin@microsoft.com
[3] The Chinese University of Hong Kong, Shatin, Hong Kong, P.R. China
zw009@ie.cuhk.edu.hk

1 Optimal Control Governed by Evolutionary PDEs

We present the related mathematical deductions of our L-PDE model. Our learning framework is based on the following PDE-constrained optimization problem:

$$\min_{\mathbf{a}} J(\{u_k\}_{k=1}^K, \mathbf{a}), \quad s.t. \begin{cases} \dfrac{\partial u_k}{\partial t} = L(u_k, \mathbf{a}), & (x,y,t) \in Q, \\ u_k = 0, & (x,y,t) \in \Gamma, \\ u_k|_{t=0} = f_k, & (x,y) \in \Omega, \end{cases} \tag{1}$$

where

$$J(\{u_k\}_{k=1}^K, \mathbf{a}) = \frac{1}{2}\sum_{k=1}^K \int_\Omega (u_k(T_f) - \tilde{u}_k)^2 d\Omega + \frac{1}{2}\sum_{i=0}^5 \alpha_i \int_0^{T_f} a_i^2(t) dt. \tag{2}$$

The above problem can be solved if we can find the Gâteaux derivative of $J(\{u_k\}_{k=1}^K, \mathbf{a})$ with respect to $\mathbf{a}$, as in this case the optimal $\mathbf{a}$ can be computed by using the conjugate gradient.

Here we introduce the following adjoint approach to provide an efficient solution to the Gâteaux derivative.

The Lagrangian function of problem (1) is:

$$\begin{aligned} &\tilde{J}(\{u_k\}_{k=1}^K, \mathbf{a}; \{\varphi_k\}_{k=1}^K) \\ &= J(\{u_k\}_{k=1}^K, \mathbf{a}) + \sum_{k=1}^K \int_Q \varphi_k [(u_k)_t - L(u_k, \mathbf{a})] dQ, \end{aligned} \tag{3}$$

where φ_k are the adjoint functions.

To find the adjoint equations for φ_k, we perturb L with respect to u. The perturbation can be written as follows:

$$
\begin{aligned}
&L(u+\varepsilon\cdot\delta u)-L(u)\\
&=\varepsilon\cdot\left(\frac{\partial L(u)}{\partial u}(\delta u)+\frac{\partial L(u)}{\partial u_x}\frac{\partial\delta u}{\partial x}+\cdots+\frac{\partial L(u)}{\partial u_{yy}}\frac{\partial^2(\delta u)}{\partial y^2}\right)+o(\varepsilon)\\
&=\varepsilon\sum_{(p,q)\in\wp}\frac{\partial L(u)}{\partial u_{pq}}\frac{\partial^{p+q}(\delta u)}{\partial x^p\partial y^q}+o(\varepsilon)\\
&=\varepsilon\sum_{(p,q)\in\wp}\sigma_{pq}(u)\frac{\partial^{p+q}(\delta u)}{\partial x^p\partial y^q}+o(\varepsilon).
\end{aligned}
\tag{4}
$$

Then we have

$$
\begin{aligned}
\delta\tilde{J}_{u_k}&=\tilde{J}(\cdots,u_k+\varepsilon\cdot\delta u_k,\cdots)-\tilde{J}(\cdots,u_k,\cdots)\\
&=\varepsilon\int_\Omega(u_k(T_f)-\tilde{u}_k)\delta u_k(T_f)d\Omega\\
&\quad+\varepsilon\int_Q\varphi_k(\delta u_k)_t dQ-\varepsilon\int_Q\varphi_k\sum_{(p,q)\in\wp}\sigma_{pq}(u_k)\frac{\partial^{p+q}(\delta u_k)}{\partial x^p\partial y^q}dQ.
\end{aligned}
\tag{5}
$$

As the perturbation δu_k should satisfy $\delta u_k|_\Gamma=0$ and $\delta u_k|_{t=0}=0$, due to the boundary and initial conditions of u_k, integrating by parts, the integration on the boundary Γ will vanish. So we have

$$
\begin{aligned}
&\delta\tilde{J}_{u_k}\\
&=\varepsilon\int_\Omega(u_k(T_f)-\tilde{u}_k)\delta u_k(T_f)d\Omega+\varepsilon\int_\Omega(\varphi_k\cdot\delta u_k)(T_f)d\Omega\\
&\quad-\varepsilon\int_Q(\varphi_k)_t\delta u_k dQ-\varepsilon\int_Q\sum_{(p,q)\in\wp}(-1)^{(p+q)}\frac{\partial^{p+q}(\sigma_{pq}(u_k)\varphi_k)}{\partial x^p\partial y^q}\delta u_k dQ+o(\varepsilon)\\
&=\varepsilon\int_Q\Bigg[(\varphi_k+u_k(T_f)-\tilde{u}_k)\delta(t-T_f)\\
&\quad-(\varphi_k)_t-\sum_{(p,q)\in\wp}(-1)^{(p+q)}\frac{\partial^{p+q}(\sigma_{pq}(u_k)\varphi_k)}{\partial x^p\partial y^q}\Bigg]\delta u_k dQ+o(\varepsilon).
\end{aligned}
\tag{6}
$$

So the adjoint equation for φ_k is

$$
\begin{cases}
\dfrac{\partial\varphi_k}{\partial t}+\displaystyle\sum_{(p,q)\in\wp}(-1)^{(p+q)}\frac{\partial^{p+q}(\sigma_{pq}(u_k)\varphi_k)}{\partial x^p\partial y^q}=0, & (x,y,t)\in Q,\\
\varphi_k=0, & (x,y,t)\in\Gamma,\\
\varphi_k|_{t=T_f}=\tilde{u}_k-u_k(T_f), & (x,y)\in\Omega,
\end{cases}
\tag{7}
$$

in order to make $\dfrac{d\tilde{J}}{du_k}=0$.

The difference in $\tilde{J}$ caused by perturbing a_i is

$$\begin{aligned}
\delta \tilde{J}_{a_i} &= \tilde{J}(\cdots, a_i + \varepsilon \cdot \delta a_i, \cdots) - \tilde{J}(\cdots, a_i, \cdots) \\
&= \frac{1}{2}\alpha_i \int_0^{T_f} (a_i + \varepsilon \cdot \delta a_i)^2(t)dt - \frac{1}{2}\alpha_i \int_0^{T_f} a_i^2(t)dt \\
&\quad - \sum_{k=1}^{K} \int_Q \varphi_k[(a_i + \varepsilon \cdot \delta a_i)(t) - a_i(t)]\mathrm{inv}_i(u_k)dQ \\
&= \varepsilon \int_0^{T_f} (\alpha_i a_i \cdot \delta a_i)(t)dt - \varepsilon \int_0^{T_f} \left(\sum_{k=1}^{K} \int_\Omega \varphi_k \mathrm{inv}_i(u_k) d\Omega \right) \delta a_i dt.
\end{aligned} \tag{8}$$

Thus we have

$$\frac{d\tilde{J}}{da_i} = \alpha_i a_i - \sum_{k=1}^{K} \int_\Omega \varphi_k \mathrm{inv}_i(u_k) d\Omega, \quad i = 0, ..., 5. \tag{9}$$