第三节课习题

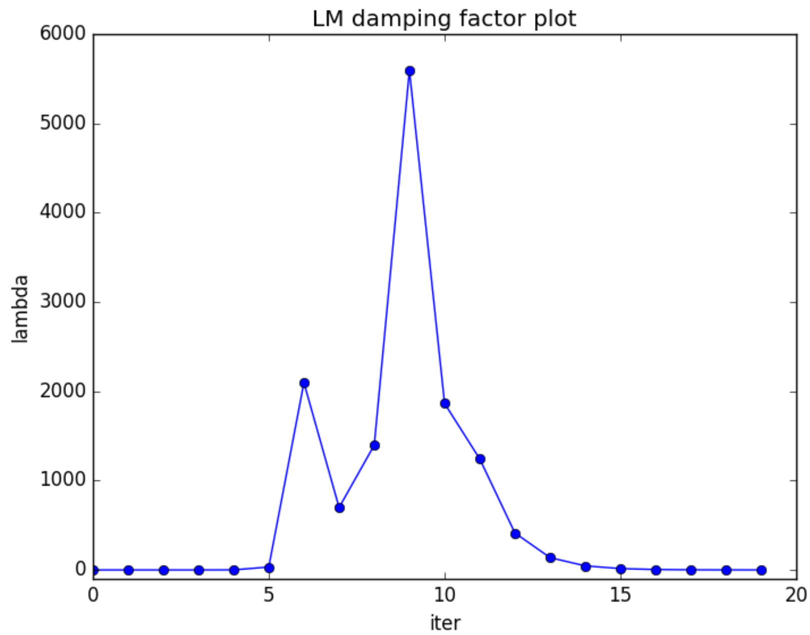
高洪臣

2019 年 6 月 30 日

第一题

样例代码给出了使用 LM 算法来估计曲线 $y = \exp(ax^2 + bx + c)$ 参数 a, b, c 的完整过程。

1. 请绘制样例代码中 LM 阻尼因子 μ 随着迭代变化的曲线图（二次修改）



2. 将曲线函数改成 $y = ax^2 + bx + c$，请修改样例代码中残差计算，雅克比计算等函数，完成曲线参数估计

   (a) `main` 函数中修改 for 循环中的观测方程

```
double y = a*x*x + b*x + c + n;
```

   (b) 修改残差计算函数为

```
virtual void ComputeResidual() override
{
  // 估计的参数
  Vec3 abc = verticies_[0]->Parameters();
  // 构建残差
  residual_(0) = (abc(0)*x_*x_ + abc(1)*x_ + abc(2)) - y_;
}
```

(c) 修改雅克比计算函数为

```
virtual void ComputeJacobians() override
{
  // 误差为1维，状态量 3 个，所以是 1x3 的雅克比矩阵
  Eigen::Matrix<double, 1, 3> jaco_abc;
  jaco_abc << x_*x_, x_ , 1;
  jacobians_[0] = jaco_abc;
}
```

(d) 曲线参数估计结果参数真实值为 $a = 1.0, b = 2.0, c = 1.0$，将代码中 `N` 改为 1000，得到参数估计值 $a = 0.999588, b = 2.0063, c = 0.968786$

3. 实现其他阻尼因子更新策略

原始程序代码实现的是 **Nielsen 策略**，现在实现 **Marquardt 策略** [1]，其算法如下：

Algorithm 1 damping strategy by Marquardt

```
if ϱ < 0.25 then
    μ := μ * 2
else if ϱ > 0.75 then
    μ := μ/3
end if
```

根据算法 3 修改`Problem::IsGoodStepInLM` 中相关代码如下

```
if(rho < 0.25) {
  currentLambda_ *= 2;
} else if (rho > 0.75) {
  currentLambda_ *= 0.3;
}

if(rho > 0) // step acceptable
  return true;
else
  return false;
```

第二题

公式推导, 根据课程知识, 完成 F, G 中如下两项的推导过程:

$$\mathbf{f}_{15} = \frac{\partial \boldsymbol{\alpha}_{b_i b_{k+1}}}{\partial \delta \mathbf{b}_k^g} = -\frac{1}{4}\left(\mathbf{R}_{b_i b_{k+1}}\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right]_\times \delta t^2\right)(-\delta t)$$

$$\mathbf{g}_{12} = \frac{\partial \boldsymbol{\alpha}_{b_i b_{k+1}}}{\partial \mathbf{n}_k^g} = -\frac{1}{4}\left(\mathbf{R}_{b_i b_{k+1}}\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right] \times \delta t^2\right)\left(\frac{1}{2}\delta t\right)$$

答:

$$
\begin{aligned}
\mathbf{f}_{15} &= \frac{\partial \boldsymbol{\alpha}_{b_i b_{k+1}}}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{\partial(\boldsymbol{\alpha}_{b_i b_k} + \boldsymbol{\beta}_{b_i b_k}\delta t + \frac{1}{2}\mathbf{a}\delta t^2)}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{2}\frac{\partial \mathbf{a}\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial\left(\mathbf{q}_{b_i b_k}\left(\mathbf{a}^{b_k} - \mathbf{b}_k^a\right) + \mathbf{q}_{b_i b_{k+1}}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{q}_{b_i b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}\omega \delta t \end{bmatrix}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{q}_{b_i b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}\omega \delta t \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -\frac{1}{2}\delta \mathbf{b}_k^g \delta t \end{bmatrix}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \exp\left(\left[-\delta \mathbf{b}_k^g \delta t\right]_\times\right)\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left(\mathbf{I} + \left[-\delta \mathbf{b}_k^g \delta t\right]_\times\right)\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left[-\delta \mathbf{b}_k^g \delta t\right]_\times\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \delta \mathbf{b}_k^g} \\
&= -\frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left(\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2\right]_\times\right)\left(-\delta \mathbf{b}_k^g \delta t\right)}{\partial \delta \mathbf{b}_k^g} \\
&= -\frac{1}{4}\mathbf{R}_{b_i b_{k+1}}\left(\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right]_\times \delta t^2\right)\left(-\delta t\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{g}_{12} &= \frac{\partial \boldsymbol{\alpha}_{b_i b_{k+1}}}{\partial \mathbf{n}_k^g} \\
&= \frac{\partial(\boldsymbol{\alpha}_{b_i b_k} + \boldsymbol{\beta}_{b_i b_k}\delta t + \frac{1}{2}\mathbf{a}\delta t^2)}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{2}\frac{\partial \mathbf{a}\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial\left(\mathbf{q}_{b_i b_k}\left(\mathbf{a}^{b_k} - \mathbf{b}_k^a\right) + \mathbf{q}_{b_i b_{k+1}}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{q}_{b_i b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}\omega \delta t \end{bmatrix}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{q}_{b_i b_k} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}\omega \delta t \end{bmatrix} \otimes \begin{bmatrix} 1 \\ \frac{1}{2}\left(\frac{1}{2}\mathbf{n}_k^g\right)\delta t \end{bmatrix}\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \exp\left(\left[\frac{1}{2}\mathbf{n}_k^g \delta t\right]_\times\right)\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left(\mathbf{I} + \left[\frac{1}{2}\mathbf{n}_k^g \delta t\right]_\times\right)\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= \frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left[\frac{1}{2}\mathbf{n}_k^g \delta t\right]_\times\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2}{\partial \mathbf{n}_k^g} \\
&= -\frac{1}{4}\frac{\partial \mathbf{R}_{b_i b_{k+1}} \left(\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\delta t^2\right]_\times\right)\left(\frac{1}{2}\mathbf{n}_k^g \delta t\right)}{\partial \mathbf{n}_k^g} \\
&= -\frac{1}{4}\mathbf{R}_{b_i b_{k+1}}\left(\left[\left(\mathbf{a}^{b_{k+1}} - \mathbf{b}_k^a\right)\right]_\times \delta t^2\right)\left(\frac{1}{2}\delta t\right)
\end{aligned}
$$

第三题

证明 ppt 中式 (9)

$$\Delta \mathbf{x}_{\mathrm{lm}} = -\sum_{j=1}^{n} \frac{\mathbf{v}_j^T \mathbf{F}'^T}{\lambda_j + \mu} \mathbf{v}_j$$

答:

已知

$$\left(\mathbf{J}^\top \mathbf{J} + \mu \mathbf{I}\right) \Delta \mathbf{x}_{\mathrm{lm}} = -\mathbf{J}^\top \mathbf{f} \quad \text{with} \quad \mu \geq 0$$

$$\mathbf{F}'(\mathbf{x}) = \left(\mathbf{J}^\top \mathbf{f}\right)^\top$$

SVD 分解

$$\mathbf{J}^\top \mathbf{J} = \mathbf{V}\mathbf{D}\mathbf{V}^\top$$

其中

$$\mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \quad s.t. \quad \lambda_j \geq 0, \quad \mathbf{V}\mathbf{V}^\top = \mathbf{V}^\top \mathbf{V} = \mathbf{I}$$

则

$$\left(\mathbf{V}\mathbf{D}\mathbf{V}^\top + \mu \mathbf{I}\right) \Delta \mathbf{x}_{\mathrm{lm}} = -\mathbf{F}'^\top$$

矩阵变换

$$\left(\mathbf{D}\mathbf{V}^\top + \mu \mathbf{V}^\top\right) \Delta \mathbf{x}_{\mathrm{lm}} = -\mathbf{V}^\top \mathbf{F}'^\top$$

$$\left(\mathbf{D}\mathbf{V}^\top + \mu \mathbf{V}^\top\right) \left(\mathbf{V}\mathbf{V}^\top\right) \Delta \mathbf{x}_{\mathrm{lm}} = -\mathbf{V}^\top \mathbf{F}'^\top$$

$$\left(\mathbf{D} + \mu \mathbf{I}\right) \mathbf{V}^\top \Delta \mathbf{x}_{\mathrm{lm}} = -\mathbf{V}^\top \mathbf{F}'^\top$$

令

$$\mathbf{D}' = \mathbf{D} + \mu \mathbf{I} = \operatorname{diag}(\lambda_1 + \mu, \lambda_2 + \mu, \ldots, \lambda_n + \mu) \quad s.t. \quad \lambda_j + \mu \geq 0$$

则其伪逆 [2]

$$\mathbf{D}'^{+}_{jj} = \begin{cases} 0 & \text{if } \mathbf{D}_{jj} = 0 \\ \mathbf{D}_{jj}^{-1} & \text{otherwise.} \end{cases}$$

所以

$$
\begin{aligned}
\Delta \mathbf{x}_{\mathrm{lm}} &= -\mathbf{V}\mathbf{D'}^{+}\mathbf{V}^{\top}\mathbf{F'}^{\top} \\
&= -\left(\sum_{j=1}^{n} \mathbf{v}_j (\lambda_j + \mu)^{-1} \mathbf{v}_j^{\top}\right) \mathbf{F'}^{\top} \qquad s.t. \quad \lambda_j + \mu > 0 \\
&= -\sum_{j=1}^{n} \frac{\mathbf{v}_j^{\top}\mathbf{F'}^{\top}}{\lambda_j + \mu} \mathbf{v}_j
\end{aligned}
$$

最终，证明

$$
\Delta \mathbf{x}_{\mathrm{lm}} = -\sum_{j=1}^{n} \frac{\mathbf{v}_j^{T}\mathbf{F'}^{T}}{\lambda_j + \mu} \mathbf{v}_j \quad s.t. \quad \lambda_j + \mu > 0
$$

参考文献

[1] Kaj Madsen, Hans Bruun Nielsen, and Ole Tingleff. Methods for non-linear least squares problems. 1999.

[2] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003.