VIO-hw3

VIO

作业

1 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
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。
3. 如果有实现其他阻尼因子更新策略可加分（选做）。

2 公式推导，根据课程知识，完成 $\mathbf{F}, \mathbf{G}$ 中如下两项的推导过程：

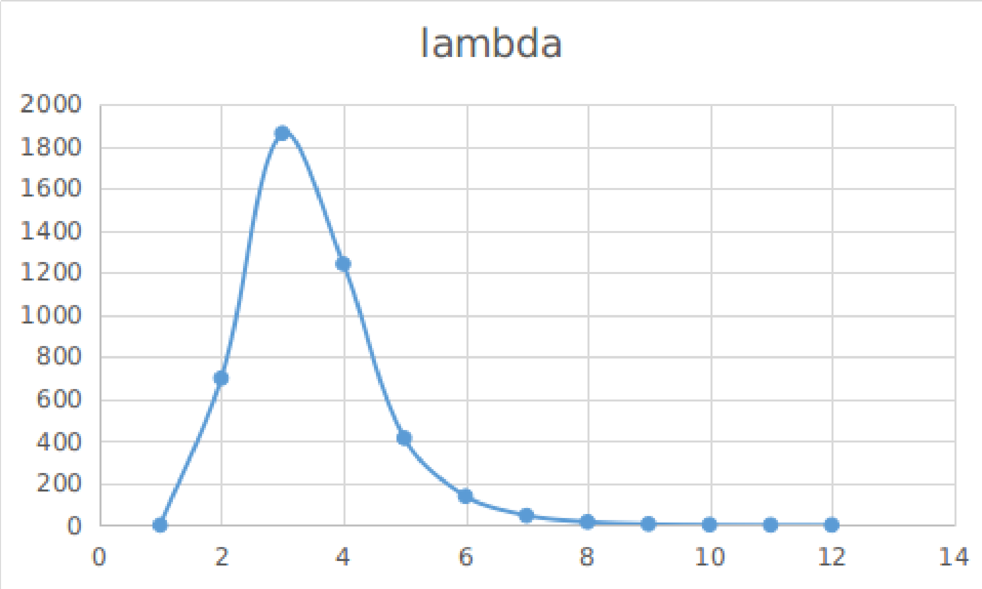
$$\mathbf{f}_{15} = \frac{\partial \boldsymbol{\alpha}_{b_i b_{k+1}}}{\partial \delta \mathbf{b}_k^g} = -\frac{1}{4}(\mathbf{R}_{b_i b_{k+1}}[(\mathbf{a}^{b_k} - \mathbf{b}_k^a)]_\times \delta t^2)(-\delta t)$$

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

3 证明式(9)。

1.LM算法

1.1 μ迭代



1.2 $y=ax^2+bx+c$

param	optimization parameters	ground truth
a	10.6107	10
b	19.6183	20
c	9.99517	10

输出：

```
10 20 10
Test CurveFitting start...
iter: 0 , chi= 61493.7 , Lambda= 0.001
iter: 1 , chi= 91.3952 , Lambda= 0.000333333
iter: 2 , chi= 91.395 , Lambda= 0.000222222
problem solve cost: 3.28693 ms
   makeHessian cost: 2.33691 ms
-------After optimization, we got these parameters :
10.6107 19.6183 9.99517
-------ground truth:
10, 20, 10
```

1.3 阻尼因子更新策略

Cui, M. , Zhao, Y. , Xu, B. , & Gao, X. W. . (2017). ***A new approach for determining damping factors in levenberg-marquardt algorithm for solving an inverse heat conduction problem.*** *International Journal of Heat and Mass Transfer, 107, 747-754.*

> 1.Initialize the weights and parameter μ, β
> 2.Stop if the number of iteration exceeds the maximum iteration or $F(w)$ is less than a desired error
> 3.After passing all training data, compute the sum of squared errors over all inputs, $F(w)$
>
> > $F(w)$即 loss function.w即正常的自变量x
>
> 4.Compute the Jacobian matrix $J(w)$
> 5.Solve Eq. (6) to obtain the weight change Δw
>
> > Eq.(6)
> >
> > $$\begin{aligned}\Delta \mathbf{w} &= -\left[\mathbf{J}(\mathbf{w})^T \mathbf{J}(\mathbf{w}) + \mu \mathbf{I}\right]^{-1} \mathbf{J}^T(\mathbf{w})\mathbf{e}(\mathbf{w}) \\ &= -[\mathbf{D}(\mathbf{w})]^{-1}\mathbf{g}(\mathbf{w})\end{aligned}$$
>
> 6.Recompute the sum of squared errors $F(w_{trial})$ using w_{trial} = w + Δw after passing all training data again, and judge
> IF $F(w_{trial})$ < $F(w)$ in step 3 THEN
>
> > $$wip = \frac{\mathrm{dot}(\mathbf{w}_{\mathrm{trial}}, \mathbf{w})}{\|\mathbf{w}_{\mathrm{trial}}\| \|\mathbf{w}\|} \qquad (-1 \le wip \le 1)$$
> >
> > $$\mu = \mu \cdot \beta^{wip} \qquad if\ F(\mathbf{w}_{\mathrm{trial}}) < F(\mathbf{w})$$
> >
> > go back to step 2

> ELSE
>
> > if $D(w)$ is positive definite
> >
> > $$\mu = \mu/\beta$$
> >
> > else
> > make $D(w)$ diagonally dominant
> >
> > $$\mathbf{D} = d_{ij}, \quad d_{ii} = \sum_{j \neq i} |d_{ij}| \quad \textbf{for all } \boldsymbol{i}$$
> >
> > $$\mu = \min(d_{ii}) \quad \textbf{for all } \boldsymbol{i}$$
> >
> > go back to step 2

2 公式推导 46/77

2.1 f_15

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

解：

$$\mathbf{f}_{15} = \frac{\partial(\alpha_{b_i b_k} + \beta_{b_i b_k}\delta t + \frac{1}{2}\mathbf{a}\delta t^2)}{\partial \delta \mathbf{b}_k^g}$$

> 其中，$(\alpha_{b_i b_k} + \beta_{b_i b_k}\delta t)$与导数无关。

$$\mathbf{f}_{15} = \frac{\partial(\frac{1}{2}\mathbf{a}\delta t^2)}{\partial \delta \mathbf{b}_k^g}$$

> 其中，$\mathbf{a}=\frac{1}{2}\left(\mathbf{q}_{b_ib_k}\left(\mathbf{a}^{b_k}-\mathbf{b}_k^a\right)+\mathbf{q}_{b_ib_{k+1}}\left(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a\right)\right)$

$$\mathbf{f}_{15}=\frac{\partial(\frac{1}{2}\ \frac{1}{2}\left(\mathbf{q}_{b_ib_k}\left(\mathbf{a}^{b_k}-\mathbf{b}_k^a\right)+\mathbf{q}_{b_ib_{k+1}}\left(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a\right)\right)\delta t^2)}{\partial\delta\mathbf{b}_k^g}$$

> 其中，$\mathbf{q}_{b_ib_k}\left(\mathbf{a}^{b_k}-\mathbf{b}_k^a\right)$ 与导数无关。

$$\mathbf{f}_{15}=\frac{1}{4}\frac{\partial\left(\mathbf{q}_{b_ib_{k+1}}\left(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a\right)\right)\delta t^2}{\partial\delta\mathbf{b}_k^g}$$

> 其中,$\mathbf{q}_{b_ib_{k+1}}=\mathbf{q}_{b_ib_k}\otimes\begin{bmatrix}1\\ \frac{1}{2}\omega\delta t\end{bmatrix}$
> $\omega=\frac{1}{2}\left(\left(\omega^{b_k}-\mathbf{b}_k^g\right)+\left(\omega^{b_{k+1}}-\mathbf{b}_k^g\right)\right)$.

$$\mathbf{f}_{15}=\frac{1}{4}\frac{\partial\left(\mathbf{q}_{b_ib_{k+1}}\left(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a\right)\right)\delta t^2}{\partial\delta\mathbf{b}_k^g}$$

$$\mathbf{f}_{15}=\frac{1}{4}\frac{\partial\mathbf{q}b_ib_k\otimes\begin{bmatrix}1\\ \frac{1}{2}\omega\delta t\end{bmatrix}\otimes\begin{bmatrix}1\\ -\frac{1}{2}\delta\mathbf{b}_k^g\delta t\end{bmatrix}\left(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a\right)\delta t^2}{\partial\delta\mathbf{b}_k^g}$$

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

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

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

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

2.2 g_12

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

$$\mathbf{g}_{12}=\frac{\partial(\alpha_{b_ib_k}+\beta_{b_ib_k}\delta t+\frac{1}{2}\mathbf{a}\delta t^2)}{\partial \delta \mathbf{n}_k^g}$$

> 部分公式代换同上(其实是敲起来好麻烦... -_-|)

$$\mathbf{g}_{12}=\frac{1}{4}\frac{\partial \mathbf{q} b_i b_k\otimes\begin{bmatrix}1\\ \frac{1}{2}\omega\delta t\end{bmatrix}\otimes\begin{bmatrix}1\\ \frac{1}{4}\delta\mathbf{n}_k^g\delta t\end{bmatrix}(\mathbf{a}^{b_{k+1}}-\mathbf{b}_k^a)\delta t^2}{\partial \delta\mathbf{n}_k^g}$$

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

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

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

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

3. 证明式9

> 需要注意以下几点：
>
> - $$\mathbf{f}(\mathbf{x})=\begin{bmatrix} f_1(\mathbf{x}) \\ \dots \\ f_m(\mathbf{x}) \end{bmatrix}_{m*1}$$
>
> $$\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}=\mathbf{J}=\begin{bmatrix} \mathbf{J}_1(\mathbf{x}) \\ \dots \\ \mathbf{J}_m(\mathbf{x}) \end{bmatrix}_{m*1}$$
>
> $\mathbf{f}(\mathbf{x})$,$\mathbf{J}$是一个$m*1$的列向量
>   所以 $F'(\mathbf{x})=\left(\mathbf{J}^\top \mathbf{f}\right)^\top=\left(\mathbf{J}^\top \mathbf{f}\right)^\top_{1*m*m*1=1*1}$ 是一个标量(所以在最后的答案里面可以放在分子的左边或者右边都是可以的)。-------------------式（1）
> $\mathbf{J}^\top\mathbf{J}=\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$ -------式(2)

推导如下:

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

$$(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top+\mu\mathbf{I})\Delta\mathbf{x}_{\text{lm}}=-\mathbf{J}^\top\mathbf{f}$$

$$(2)=>(\mathbf{V}(\mathbf{\Lambda}+\mu\mathbf{I})\mathbf{V}^\top)\Delta\mathbf{x}_{\text{lm}}=-\mathbf{J}^\top\mathbf{f}$$

$$(\mathbf{\Lambda}+\mu\mathbf{I})\mathbf{V}^\top\Delta\mathbf{x}_{\text{lm}}=-\mathbf{V}^\top\mathbf{J}^\top\mathbf{f}$$

$$(1)=>(\mathbf{\Lambda}+\mu\mathbf{I})\mathbf{V}^\top\Delta\mathbf{x}_{\text{lm}}=-\mathbf{V}^\top F'(\mathbf{x})^\top$$

$$\mathbf{V}^\top\Delta\mathbf{x}_{\text{lm}}=-(\mathbf{\Lambda}+\mu\mathbf{I})^{-1}\mathbf{V}^\top F'(\mathbf{x})^\top$$

$$\mathbf{V}^\top \Delta\mathbf{x}_{\text{lm}} = -\begin{bmatrix} \dfrac{1}{\lambda_1+\mu} & 0 & 0 & \dots \\ 0 & \dfrac{1}{\lambda_2+\mu} & 0 & \dots \\ & \dots & \dfrac{1}{\lambda_j+\mu} \end{bmatrix} \mathbf{V}^\top F'(\mathbf{x})^\top$$

$$\Delta\mathbf{x}_{\text{lm}} = -[\,\mathbf{v}_1 \dots \mathbf{v}_2\,] \begin{bmatrix} \dfrac{1}{\lambda_1+\mu} & 0 & 0 & \dots \\ 0 & \dfrac{1}{\lambda_2+\mu} & 0 & \dots \\ & \dots & \dfrac{1}{\lambda_j+\mu} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \dots \\ \mathbf{v}_2 \end{bmatrix} F'(\mathbf{x})^\top$$

$$\Delta\mathbf{x}_{\text{lm}} = -\sum_{j=1}^{n} \frac{\mathbf{v}_j^\top \mathbf{F}'^\top}{\lambda_j+\mu}\,\mathbf{v}_j$$