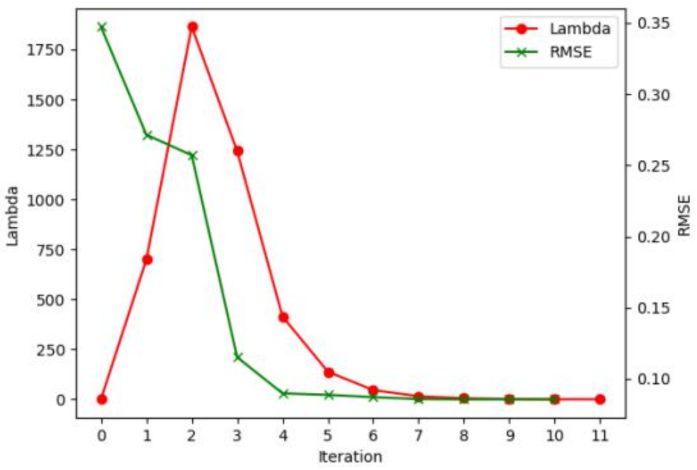
深蓝学院

shenlanxueyuan.com

第三章作业分享

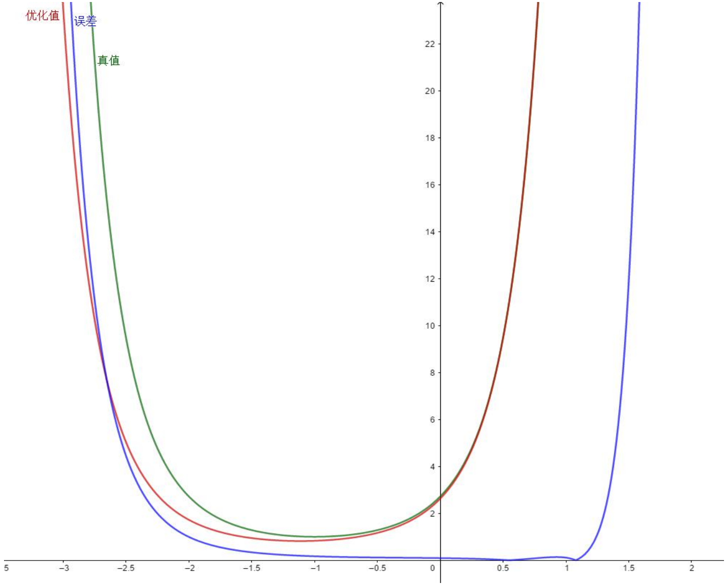
第一题

● 1.1 绘制阻尼因子随迭代的变化曲线图

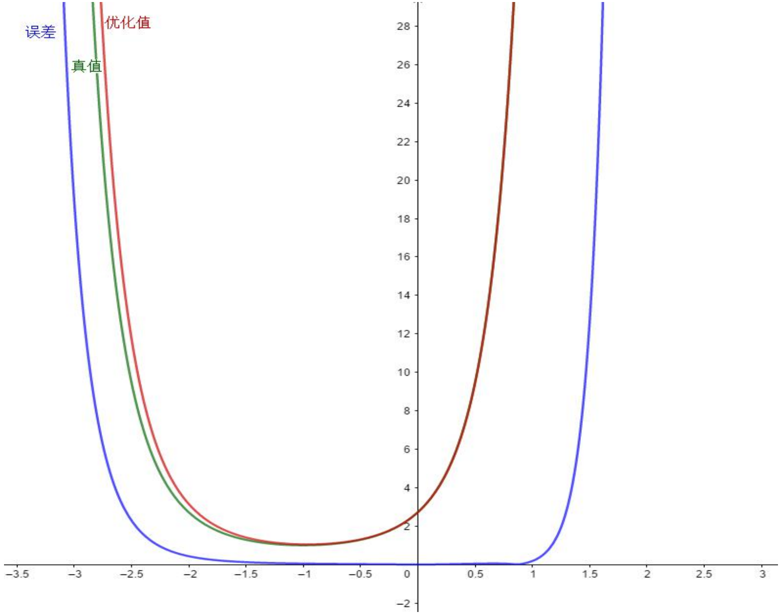


第一题

● 1.1 绘制阻尼因子随迭代的变化曲线图



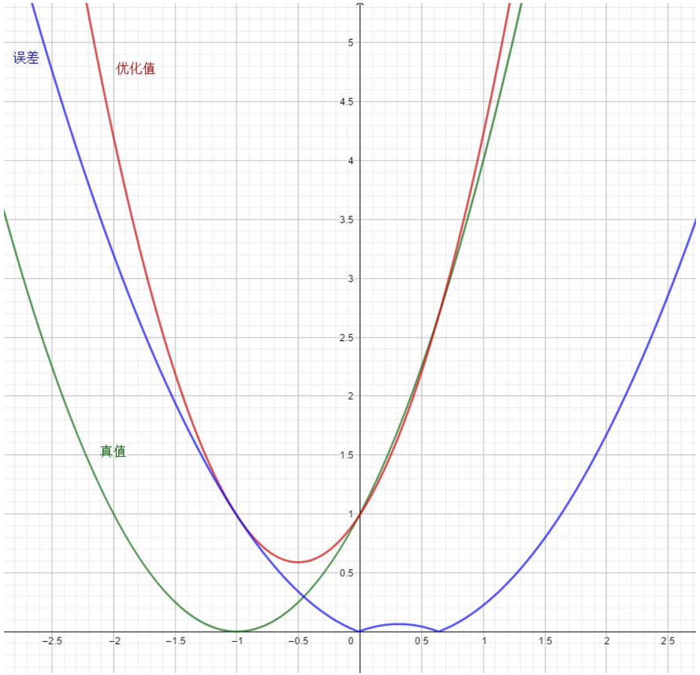
默认采样区间：[0, 1]



扩大后的采样区间：[-1, 1]

第一题

● 1.2 LM求解$y = ax^2 + bx + c$




```
        // 观测 y
        // double y = std::exp( a*x*x + b*x + c ) + n;
        double y = a*x*x + b*x + c + n;
//        double y = std::exp( a*x*x + b*x + c );
```

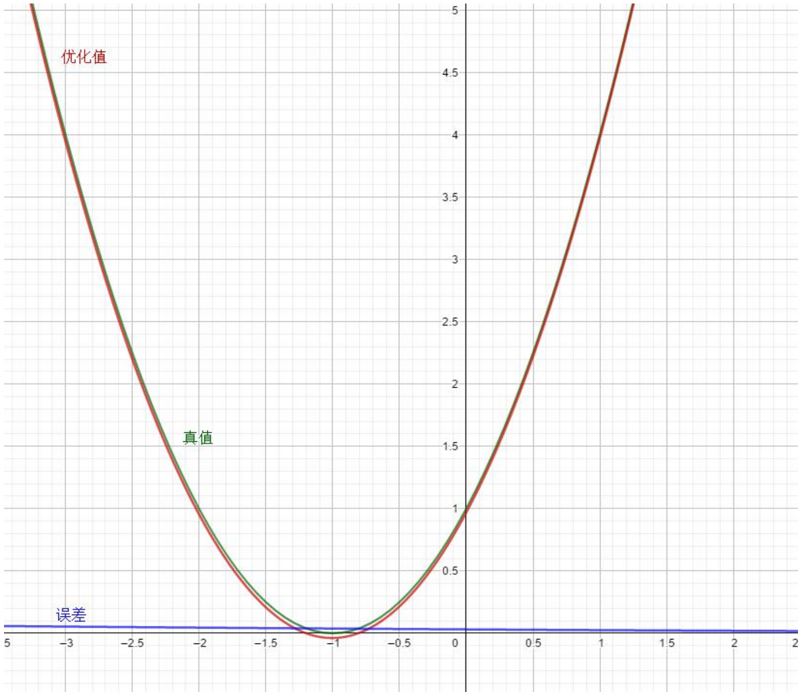
```
// 计算残差对变量的雅克比
virtual void ComputeJacobians() override
{
    Vec3 abc = verticies_[0]->Parameters();
    double exp_y = std::exp( x: abc( index: 0)*x_*x_ + abc( index: 1)*x_ + abc( index: 2) );

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

```
// 计算曲线模型误差
virtual void ComputeResidual() override
{
    Vec3 abc = verticies_[0]->Parameters();  // 估计的参数
    // residual_(0) = std::exp( abc(0)*x_*x_ + abc(1)*x_ + abc(2) ) - y_;  // 构建残差
    residual_( index: 0) = abc( index: 0)*x_*x_ + abc( index: 1)*x_ + abc( index: 2) - y_;  // 构建残差
}
```

第一题

- 1.2 LM求解$y = ax^2 + bx + c$



采样区间：[0, 10]

第一题

●1.3 实现更优秀的阻尼因子策略

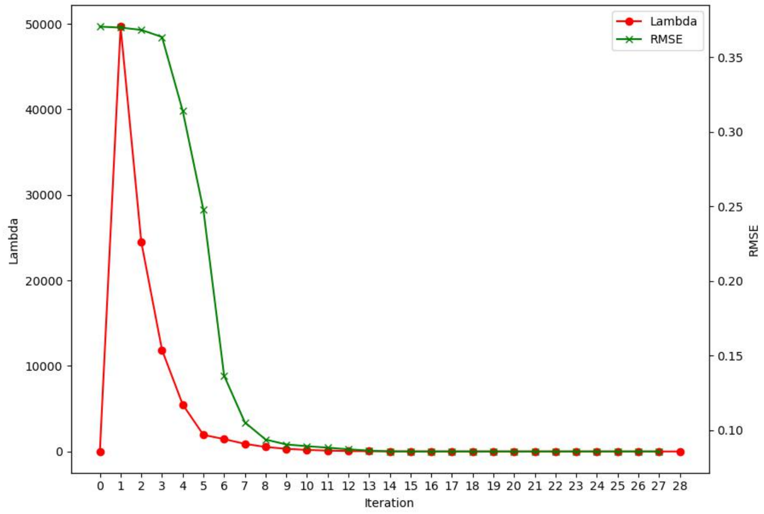
论文 The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems 4.1.1 节中提供了 3 种不同的阻尼因子更新策略。我的作业里主要实现了第2种。

策略摘要：

1. 选择一个合适的系数τ（由经验确定），并设置初始阻尼因子$\lambda_0 = \tau \max\{\mathrm{diag}[\boldsymbol{J}^\top \boldsymbol{W} \boldsymbol{J}]\}$（$\boldsymbol{W}$为权重矩阵）。
2. 计算$\alpha = \left(\left(\boldsymbol{J}^\top \boldsymbol{W}(\boldsymbol{y}-\hat{\boldsymbol{y}}(\boldsymbol{p}))\right)^\top \boldsymbol{h}\right) \Big/ \left(\left(\chi^2(\boldsymbol{p}+\boldsymbol{h})-\chi^2(\boldsymbol{p})\right)/2 + 2\left(\boldsymbol{J}^\top \boldsymbol{W}(\boldsymbol{y}-\hat{\boldsymbol{y}}(\boldsymbol{p}))\right)^\top \boldsymbol{h}\right)$
3. 如果$\rho(\alpha \boldsymbol{h})$大于阈值（经验确定，一般可设置为0），更新$\lambda \leftarrow \max\{\lambda/(1+\alpha), 10^{-7}\}$，更新$\boldsymbol{p} \leftarrow \boldsymbol{p} + \alpha \boldsymbol{h}$
4. 否则，更新$\boldsymbol{\lambda} \leftarrow \boldsymbol{\lambda} + |\chi^2(\boldsymbol{p}+\alpha \boldsymbol{h})-\chi^2(\boldsymbol{p})|/2\alpha$

第一题

●1.3 实现更优秀的阻尼因子策略



同样的初始参数拟合e^{ax^2+bx+c}

迭代数比Nielson策略多，RMSE达到同样数量级的迭代数量接近。

最终的RMSE接近，拟合得到的参数效果近似（区间内都能较好逼近，区间外偏差较大）

第二题

- $\boldsymbol{f}_{15}$

$$\boldsymbol{\alpha}_{b_ib_{k+1}} = \boldsymbol{\alpha}_{b_ib_k} + \boldsymbol{\beta}_{b_ib_k}\delta t + \frac{1}{2}\boldsymbol{a}\delta t^2$$

$$\begin{aligned}\boldsymbol{a}\delta t^2 &= \frac{1}{2}\left(\boldsymbol{q}_{b_ib_k}\left(\boldsymbol{a}^{b_k} - \boldsymbol{b}_k^a\right) + \boldsymbol{q}_{b_ib_{k+1}}\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\right)\delta t^2 \\ &= \frac{1}{2}\left(\boldsymbol{q}_{b_ib_k}\left(\boldsymbol{a}^{b_k} - \boldsymbol{b}_k^a\right) + \left(\boldsymbol{q}_{b_ib_k} \otimes \begin{pmatrix} 1 \\ \frac{1}{2}\boldsymbol{\omega}\delta t \end{pmatrix}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2\right)\end{aligned}$$

$$\boldsymbol{\omega} = \frac{1}{2}\left(\left(\boldsymbol{\omega}^{b_k} - \boldsymbol{b}_k^g\right) + \left(\boldsymbol{\omega}^{b_{k+1}} - \boldsymbol{b}_k^g\right)\right) = \frac{1}{2}\left(\boldsymbol{\omega}^{b_k} + \boldsymbol{\omega}^{b_{k+1}}\right) - \boldsymbol{b}_k^g$$

$$\begin{aligned}\frac{\partial \boldsymbol{\alpha}_{b_ib_{k+1}}}{\partial \delta \boldsymbol{b}_k^g} &= \frac{\partial \frac{1}{2}\cdot\frac{1}{2}\left(\boldsymbol{q}_{b_ib_k} \otimes \begin{pmatrix} 1 \\ \frac{1}{2}\boldsymbol{\omega}\delta t \end{pmatrix}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &= \frac{\partial \frac{1}{4}\left(\boldsymbol{q}_{b_ib_k} \otimes \begin{pmatrix} 1 \\ \frac{1}{2}\left(\boldsymbol{\omega} - \delta\boldsymbol{b}_k^g\right)\delta t \end{pmatrix}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &= \frac{\partial \frac{1}{4}\left(\boldsymbol{R}_{b_ib_k} e^{\left[\left(\boldsymbol{\omega} - \delta\boldsymbol{b}_k^g\right)\delta t\right]_\times}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &\approx \frac{\partial \frac{1}{4}\left(\boldsymbol{R}_{b_ib_k} e^{[\boldsymbol{\omega}\delta t]_\times} e^{\left[-\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\delta\boldsymbol{b}_k^g\delta t\right]_\times}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &\approx \frac{\partial \frac{1}{4}\left(\boldsymbol{R}_{b_ib_{k+1}}\left(\boldsymbol{I} - \left[\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\delta\boldsymbol{b}_k^g\delta t\right]_\times\right)\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &= \frac{\partial -\frac{1}{4}\boldsymbol{R}_{b_ib_{k+1}}\left[-\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\delta\boldsymbol{b}_k^g\delta t\right]_\times\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &= \frac{\partial -\frac{1}{4}\boldsymbol{R}_{b_ib_{k+1}}\left[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right]_\times\left(-\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\delta\boldsymbol{b}_k^g\delta t\right)\delta t^2}{\partial \delta \boldsymbol{b}_k^g} \\ &= -\frac{1}{4}\boldsymbol{R}_{b_ib_{k+1}}\left[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right]_\times\left(-\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\delta t\right)\delta t^2 \\ &\approx -\frac{1}{4}\left(\boldsymbol{R}_{b_ib_{k+1}}\left[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right]_\times\right)\delta t^2(-\delta t)\end{aligned}$$

第二题

- $\boldsymbol{g}_{12}$

$$\boldsymbol{\alpha}_{b_ib_{k+1}} = \boldsymbol{\alpha}_{b_ib_k} + \boldsymbol{\beta}_{b_ib_k}\delta t + \frac{1}{2}\boldsymbol{a}\delta t^2$$

$$\begin{aligned}\boldsymbol{a}\delta t^2 &= \frac{1}{2}\left(\boldsymbol{q}_{b_ib_k}\left(\boldsymbol{a}^{b_k} - \boldsymbol{b}_k^a\right) + \boldsymbol{q}_{b_ib_{k+1}}\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\right)\delta t^2 \\ &= \frac{1}{2}\left(\boldsymbol{q}_{b_ib_k}\left(\boldsymbol{a}^{b_k} - \boldsymbol{b}_k^a\right) + \left(\boldsymbol{q}_{b_ib_k} \otimes \begin{pmatrix} 1 \\ \frac{1}{2}\boldsymbol{\omega}\delta t \end{pmatrix}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2\right)\end{aligned}$$

$$\begin{aligned}\boldsymbol{\omega} &= \frac{1}{2}\left(\left(\left(\boldsymbol{\omega}^{b_k} + \boldsymbol{n}_k^g\right) - \boldsymbol{b}_k^g\right) + \left(\left(\boldsymbol{\omega}^{b_{k+1}} + \boldsymbol{n}_{k+1}^g\right) - \boldsymbol{b}_k^g\right)\right) \\ &= \frac{1}{2}\left(\boldsymbol{\omega}^{b_k} + \boldsymbol{\omega}^{b_{k+1}}\right) - \boldsymbol{b}_k^g + \frac{1}{2}\boldsymbol{n}_{k+1}^g + \frac{1}{2}\boldsymbol{n}_k^g\end{aligned}$$

$$\begin{aligned}\frac{\partial \boldsymbol{\alpha}_{b_ib_{k+1}}}{\partial \delta \boldsymbol{n}_{\boldsymbol{b}_k}^g} &= \frac{\partial \frac{1}{2}\cdot\frac{1}{2}\left(\boldsymbol{q}_{b_ib_k} \otimes \begin{pmatrix} 1 \\ \frac{1}{2}\left(\boldsymbol{\omega} + \frac{1}{2}\boldsymbol{n}_{\boldsymbol{b}_k^g}\right)\delta t \end{pmatrix}\right)\left(\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a\right)\delta t^2}{\partial \delta \boldsymbol{n}_{\boldsymbol{b}_k}^g} \\ &= \frac{\partial -\frac{1}{4}\boldsymbol{R}_{b_ib_{k+1}}[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a]_\times\left(\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\frac{1}{2}\boldsymbol{n}_{\boldsymbol{b}_k^g}\delta t\right)\delta t^2}{\partial \delta \boldsymbol{n}_{\boldsymbol{b}_k}^g} \\ &= -\frac{1}{4}\boldsymbol{R}_{b_ib_{k+1}}[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a]_\times\left(\boldsymbol{J}_r(\boldsymbol{\omega}\delta t)\frac{1}{2}\delta t\right)\delta t^2 \\ &\approx -\frac{1}{4}\left(\boldsymbol{R}_{b_ib_{k+1}}[\boldsymbol{a}^{b_{k+1}} - \boldsymbol{b}_k^a]_\times\right)\delta t^2\left(\frac{1}{2}\delta t\right)\end{aligned}$$

第三题

●证明式(9)

1. 所有损失函数f_i都是实函数+$\boldsymbol{J}^\top\boldsymbol{J}$是半正定矩阵 =>$\boldsymbol{J}^\top\boldsymbol{J}$是实对称矩阵，可对角正交化

2. 对角正交化形式：$\boldsymbol{J}^\top\boldsymbol{J}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^\top$ => $\boldsymbol{V}$是正交矩阵，$\boldsymbol{V}$满秩，colspan $\boldsymbol{V}=\mathbb{R}^{\boldsymbol{n}}$

3. $\boldsymbol{\Delta x}_{\mathrm{lm}}, -\boldsymbol{F}'^\top$ 可被$\{\boldsymbol{v_1},\boldsymbol{v_2},\ldots,\boldsymbol{v}_n\}$线性表示。$\{\boldsymbol{v_1},\boldsymbol{v_2},\ldots,\boldsymbol{v}_n\}$是$\boldsymbol{J}^\top\boldsymbol{J}$的特征向量，对应特征值$\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$。

$$
\begin{aligned}
(\boldsymbol{J}^\top\boldsymbol{J}+\mu\boldsymbol{I})\Delta\boldsymbol{x}_{\mathrm{lm}} &= \boldsymbol{J}^\top\boldsymbol{J}\Delta\boldsymbol{x}_{\mathrm{lm}}+\mu\boldsymbol{I}\Delta\boldsymbol{x}_{\mathrm{lm}} \\
&= \boldsymbol{J}^\top\boldsymbol{J}\sum_{i=1}^{n} x_i\boldsymbol{v_i}+\mu\sum_{i=1}^{n} x_i\boldsymbol{v_i} \\
&= \sum_{i=1}^{n} x_i\boldsymbol{J}^\top\boldsymbol{J}\boldsymbol{v_i}+\mu x_i\boldsymbol{v_i} \\
&= \sum_{i=1}^{n} x_i\lambda_i\boldsymbol{v_i}+\mu x_i\boldsymbol{v_i} \\
&= \sum_{i=1}^{n} x_i(\lambda_i+\mu)\boldsymbol{v_i}
\end{aligned}
$$

$$
\begin{aligned}
-\boldsymbol{F}'^\top = -\sum_{i=1}^{n} f_i\boldsymbol{v_i} &= \sum_{i=1}^{n} x_i(\lambda_i+\mu)\boldsymbol{v_i} \\
x_i(\lambda_i+\mu) &= -f_i \\
x_i &= -\frac{f_i}{\lambda_i+\mu}
\end{aligned}
$$

$$
\begin{aligned}
f_i\boldsymbol{v_i} &= \mathrm{proj}_{\boldsymbol{v_i}}\boldsymbol{F}'^\top = \frac{\boldsymbol{F}'^\top\cdot\boldsymbol{v_i}}{\boldsymbol{v_i}\cdot\boldsymbol{v_i}}\boldsymbol{v_i} \\
f_i\boldsymbol{v_i} &= \left(\boldsymbol{F}'^\top\cdot\boldsymbol{v_i}\right)\boldsymbol{v_i} \\
f_i &= \boldsymbol{F}'^\top\cdot\boldsymbol{v_i}
\end{aligned}
$$

$$
\boldsymbol{\Delta x}_{lm} = -\sum_{i=1}^{n}\frac{\boldsymbol{F}'^\top\cdot\boldsymbol{v}_i}{\lambda_i+\mu}\boldsymbol{v}_i = -\sum_{i=1}^{n}\frac{\boldsymbol{v}_i^\top\boldsymbol{F}'^\top}{\lambda_i+\mu}\boldsymbol{v}_i
$$

感谢各位聆听！

Thanks for Listening