Solve Nonlinear Least-Squares Problem with the Gauss-Newton Methods.

What is linear/nonlinear?

- Linear: A polynomial of degree 1.
- Nonlinear: A function cannot be expressed as a polynomial of degree 1.

What is Least-Squares Problem?

f(x) is the objective function, x is the parameter vector.
The least-squares problem tries to find optimal parameters to minimize the overall cost.

$$cost = \sum_{i=0}^{n} f(x)^2 \tag{1}$$

What is Gauss-Newton Methods?

The gauss–newton methods is used to solve nonlinear least-squares problems.
Unlike Newton's method, gauss-newton methods are not necessary to calculate second derivatves, which may difficult to compute in some cases.
Gauss–newton methods update x using iterative method, the update amount is Δx.

$$\Delta x = -H^{-1} g \tag{2}$$

here: g is the gradient vector; H is the hessian matrix.

$$g = J^T e \tag{3}$$

$$H \approx J^T J \tag{4}$$

J is the jacobian matrix, e is the residual vector.

The jacabian matrix of 3d rotation or 3d transform

If we define the increment of SO3/SE3 as:

$$T(x_0 \oplus \delta) \triangleq T(x_0) \exp(\delta) \tag{5}$$

The $\delta \in \mathfrak{so}(3)$ or $\delta \in \mathfrak{se}(3)$

We use a first-order Taylor expansion to approximate the original equation:

$$T(x_0 \oplus \delta) = T_0 \exp(\delta) \cong T_0 + T_0\widehat{\delta} \tag{6}$$

The f(x) is the objective function.
We want to find the optimal parameters (x) that minimize the result of the objective function.

$$f(x) = T(x)a - b \tag{7}$$

a is the target point:
b is the reference point.

We can use gauss-newton method to solve this problem.
According to gauss-newton method, we need to find the Jacobian matrix
for f(x).

$$\begin{aligned} \dot{f} = \frac{T_0 \exp\left(\delta\right) a - T_0 a}{\delta} \\ \cong \frac{T_0 a + T_0\widehat{\delta} a - T_0 a}{\delta} \\ = \frac{T_0\widehat{\delta} a}{\delta} \\ = -\frac{T_0 \delta \widehat{a}}{\delta} \\ = -T_0\widehat{a} \end{aligned} \tag{8}$$

When $\delta \in \mathfrak{so}(3)$

T_0 is a 3d rotation matrix(R_0),
and $\widehat{a}$ is defined as a skew symmetric matrix for vector a

$$\dot{f} = -R_0[a]_\times \tag{9}$$

When $\delta \in \mathfrak{se}(3)$

$$\delta = [\ v, \omega] \tag{10}$$

ω: the parameters of rotation.

v: the parameters of translation.

$$\widehat{\delta} = \begin{bmatrix} [\omega]_\times & v \\ 0 & 0 \end{bmatrix} \tag{11}$$

$$
\begin{aligned}
\dot{f} &= \frac{R_0 \hat{\delta} a}{\delta} \\
&= \frac{T_0 \begin{bmatrix} [\omega]_\times & v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix}}{[\omega,\ v]} \\
&= \frac{T_0 \begin{bmatrix} [-a]_\times & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \omega \\ v \end{bmatrix}}{[\omega,\ v]} \\
&= T_0 \begin{bmatrix} [-a]_\times & I \\ 0 & 0 \end{bmatrix}
\end{aligned} \tag{12}
$$