Multiple View Geometry: Exercise Sheet 10

Prof. Dr. Daniel Cremers,
Mohammed Brahimi, Zhenzhang Ye, Regine Hartwig
Computer Vision Group, TU Munich
`Wednesdays 16:15–17:45 at Hörsaal 2, "Interims I" (5620.01.102), and on RBG Live`

Exercise: July 17th, 2024

---

1. **Gauss-Newton Method** When optimizing a function $F(x) = \frac{1}{2}\|r(x)\|_2^2$ with residual $r(x)$, the Gauss-Newton method approximates the residual using a Taylor expansion:

$$r(x_0 + \Delta x) \approx r(x_0) + J_r(x_0)\Delta x \tag{1}$$

The minimization problem thus is

$$\min_{\Delta x} \frac{1}{2}\|r_0 + J_0\Delta x\|_2^2 \tag{2}$$

with a slight abuse of notation $J := J_r(x_0)$ and $r_0 := r(x_0)$.

   (a) Compute the gradient of $\frac{1}{2}\|r_0 + J\Delta x\|_2^2$ w.r.t. Δx.
   (b) Solve the optimality condition for Δx.
   (c) What problems can occur when solving for Δx?

2. **Levenberg-Marquardt Method** One way to motivate the Levenberg-Marquardt method is to tackle the previously discussed problem by adding the damping term as follows:

$$\left(J^\top J + \lambda D^T D\right)\Delta x = -J^\top r. \tag{3}$$

However, this can also be seen as a regularized version of the Gauss-Newton method.

$$\min_{\Delta x} \frac{1}{2}\|r + J\Delta x\|_2^2 + \frac{\lambda}{2}\|D\Delta x\|_2^2. \tag{4}$$

   (a) Compute the gradient of the new cost function w.r.t. Δx.
   (b) Solve the optimality condition for Δx.
   (c) What is the effect of λ on the solution?

3. **Levenberg-Marquardt for Bundle Adjustment** Now, we apply the Levenberg-Marquardt method to the bundle adjustment problem. The variables are as follows:

   - n_p: number poses
   - n_l: number landmarks
   - d_p: umber of camera parameters
   - $x_p \in \mathbb{R}^{n_p d_p}$: camera parameters
   - $x_l \in \mathbb{R}^{n_l 3}$: landmark positions
   - $x = \begin{bmatrix} x_p \\ x_l \end{bmatrix}$

We resuse the results from the previous problem

$$\min_{\Delta x} \frac{1}{2}\|r + J\Delta x\|_2^2 + \frac{\lambda}{2}\|D\Delta x\|_2^2. \tag{5}$$

which is the following optimality condition

$$J^\top r_0 = \underbrace{(J^\top J + \lambda D^T D)}_{H} \Delta x. \tag{6}$$

Now we split the Jacobian and damping into two parts $J = \begin{bmatrix} J_p & J_l \end{bmatrix}$ and $D = \begin{bmatrix} D_p & D_l \end{bmatrix}$.

(a) What is the dimension of H? What problems can occur when solving for Δx? What are the dimensions of J_p, J_ℓ, D_p, D_ℓ? Lets see what we can do...

(b) Rewrite the optimality condition by rewriting the matrix H into the block matrix form, yielding the normal equation:

$$\begin{pmatrix} U & W \\ W^\top & V \end{pmatrix} \begin{pmatrix} \Delta x_p \\ \Delta x_\ell \end{pmatrix} = \begin{pmatrix} b_p \\ b_\ell \end{pmatrix}. \tag{7}$$

What are U, W, V, b_p, b_ℓ and their dimensions?

(c) The Schur complement is allowing us to first solve for Δx_p using the Schur complement S. Derive the Schur complement S and the vector $\tilde{b}$ for the reduced system: $S\Delta x_p = \tilde{b}$.

(d) What is the dimension of S?

4. **Power Bundle Adjustment** The goal of Power Bundle Adjustment is to solve the reduced system $S\Delta x_p = \tilde{b}$ efficiently.

(a) From the lecture, we know that computing the inverse of the Schur component can be approximated by a matrix power series. Specifically, we have:

$$\begin{aligned} S &= U(I - U^{-1}WV^{-1}W^\top) \\ \to S^{-1} &= (I - U^{-1}WV^{-1}W^\top)^{-1}U^{-1} \\ \to S^{-1} &\approx \sum_{i=0}^{m} (U^{-1}WV^{-1}W^\top)^i U^{-1}. \end{aligned} \tag{8}$$

To apply the matrix power series, we need to guarantee the spectral norm of the matrix is smaller than 1, i.e. show that the eigenvalue μ of $U^{-1}WV^{-1}W^\top$ satisfies $0 \leq \mu < 1$.
Hint: Consider the similar matrix $U^{-1/2}WV^{-1}WU^{-1/2}$ for $U^{-1}WV^{-1}W^\top$ and show $U^{-1/2}WV^{-1}WU^{-1/2}$ is positive semi-definite. Additionally, the similar matrix $U^{-1/2}SU^{-1/2}$ for $U^{-1}S$ and show it is positive definite.

5. **Dense RGB-D Tracking** In the previous bundle adjustment problem, we have seen how to optimize the camera parameters x_p and landmark positions x_l. In the context of direct approaches, we optimize for the extrinsic camera parameters $x_p = [\xi_1, ..., \xi_{n_p}]$ using the photometric error as a residual and frame wise depth map h provided. With known camera poses, the 3D geometry can thus be densely be reconstructed. No need to optimize for landmark positions x_l. The residual is as follows:

$$E(x_p) = \sum_i \int_{\Omega_1} \| \underbrace{I_1(x) - I_i(\Pi g_{\xi_i}(hx))}_{r_x(\xi_i)} \|^2 dx \tag{9}$$

where I_1 and I_i are the intensity images, Π is the projection operator, g_{ξ_i} is the rigid transorm depending on the camera pose. The integral is over the image domain Ω_1 with x here being the homogeneous image coordinate and h its depth in the first frame.

(a) Using the results from previous problems, state the solution for minimizing the residual $r_x(\xi_i)$ using the Levenberg-Marquardt method.

(b) Compute the Jacobian of the residual $r_x(\xi_i)$ w.r.t. the camera parameters ξ_i, but don't explicitly compute $\frac{d}{d\xi_i} g_{\xi_i}(hx)$.