Bundle Adjustment:

Residual structure = $\begin{bmatrix} f_{11} \\ g_{11} \\ f_{12} \\ g_{12} \\ \vdots \\ f_{MN} \\ g_{MN} \end{bmatrix}$

where for f_{ij}, i is for the ith camera, j is for the jth point observed in the camera

$f_{ij} = \hat{x}_{ij} - x_{ij}$

$= P_i X_{ij1} - x_{ij}$

$g_{ij} = \hat{y}_{ij} - y_{ij}$

$= P_i X_{ij2} - y_{ij}$

Here $P_i X_{ij} = \overline{X_{ij}}$

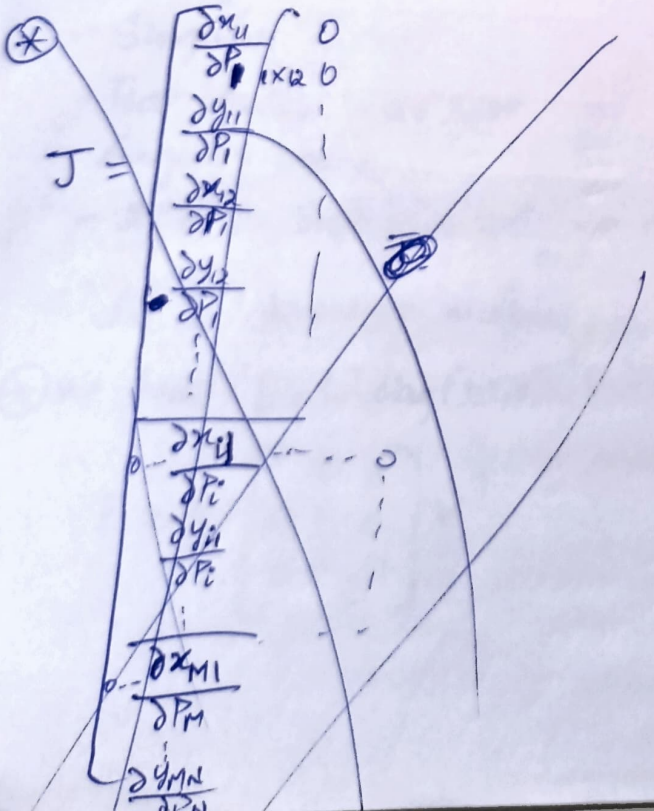
* As we can see from the residual structure, we are first keeping the camera constant and varying the points observed by that camera
* So for the Jacobian calculation, the same order is followed

$J =$

$\begin{bmatrix}
\frac{\partial x_{11}}{\partial P_1} & 0 & \cdots & \frac{\partial x_{11}}{\partial X_1} & 0 & \cdots \\
\frac{\partial y_{11}}{\partial P_1} & 0 & \cdots & \frac{\partial y_{11}}{\partial X_1} & 0 & \cdots \\
\vdots & \vdots & & 0 & \frac{\partial x_{12}}{\partial X_2} & \cdots \\
\frac{\partial x_{1N}}{\partial P_1} & 0 & \cdots & \vdots & \vdots & \\
\frac{\partial y_{1N}}{\partial P_1} & 0 & \cdots & \cdots & 0 & \frac{\partial x_{1N}}{\partial X_N} \\
\hline
0 & \frac{\partial x_{21}}{\partial P_2} & \cdots & & & \frac{\partial y_{1N}}{\partial X_N} \\
0 & \frac{\partial y_{21}}{\partial P_2} & \cdots & & & \\
\vdots & \vdots & & & & \\
\hline
0 & 0 & \cdots & \frac{\partial x_{i1}}{\partial P_i} & \cdots & 0 \\
0 & 0 & \cdots & \frac{\partial y_{i1}}{\partial P_i} & & \\
\vdots & \vdots & & \vdots & & \\
0 & 0 & \cdots & \frac{\partial x_{ij}}{\partial P_i} & 0 & \\
0 & 0 & \cdots & \frac{\partial y_{ij}}{\partial P_i} & & \\
\hline
\vdots & \vdots & & \vdots & & \\
\hline
0 & 0 & \cdots & 0 & \cdots & \frac{\partial x_{M1}}{\partial P_M} \\
 & & & & & \frac{\partial y_{M1}}{\partial P_M} \\
 & & & & & \vdots \\
 & & & & & \frac{\partial x_{MN}}{\partial P_M} \\
 & & & & & \frac{\partial y_{MN}}{\partial P_M}
\end{bmatrix}$

$2N$ points

1×12 → 12 entries of P_i matrix of camera i



⊛ In the Jacobian, we keep the order same as the residual calculation and first compute over all points for each camera.

⊛ Hence, in the Jacobian we have $\frac{\partial x_{ij}}{\partial P_i}, \frac{\partial y_{ij}}{\partial P_i}, \forall j = 1, \ldots, N$ and all other columns 0 in the first N rows

⊛ $\frac{\partial x_{ij}}{\partial P_i}$ is of dimension 1×12 because of 12 entries in P_i

⊛ Similarly, $\frac{\partial x_{ij}}{\partial X_j}$ is of dimension 1×3 as X_j has 3 coords

⊛ So, the Jacobian has shape: $12MN$ rows for the $2N$ pts (x and y) for each of the M cameras and $12M + 3N$ cols

12 params per each cam and first M vectors in a row are by cams due to our residual formulation

Then, 3 params per each point and last N vectors in the row are by points