



## Multiple View Geometry: Solution Sheet 6

Prof. Dr. Daniel Cremers, Simon Weber, Tarun Yenamandra

Computer Vision Group, TU Munich

Wednesdays 16:00–18:15 at Hörsaal 2, "Interims I"

(5620.01.102), and on RBG Live

Exercise: July 06th, 2022

### Part I: Theory

1. (a)  $E$  is essential matrix  $\Rightarrow \Sigma = \text{diag}\{\sigma, \sigma, 0\}$ :

$$R_z(\pm\frac{\pi}{2})\Sigma = \begin{pmatrix} 0 & \mp 1 & 0 \\ \pm 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mp\sigma & 0 \\ \pm\sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = -(R_z(\pm\frac{\pi}{2})\Sigma)^\top$$

$$\begin{aligned} -\hat{T}^\top &= -(UR_z\Sigma U^\top)^\top \\ &= U(-R_z\Sigma)^\top U^\top \\ &= UR_z\Sigma U^\top \\ &= \hat{T} \end{aligned}$$

- (b) Since  $U, V$  are orthogonal with determinant 1 (see lecture), they are rotation matrices. Since  $\text{SO}(3)$  is a group and thus closed under multiplication,  $R \in \text{SO}(3)$ .

*Alternative longer proof:*

- i.  $U, V$  are orthogonal matrices  $\Rightarrow U^\top U = \mathbb{1}$  and  $VV^\top = \mathbb{1}$

$$R_z \text{ is a rotation matrix } \Rightarrow R_z R_z^\top = \mathbb{1}$$

$$\begin{aligned} R^\top R &= (UR_z^\top V^\top)^\top (UR_z^\top V^\top) \\ &= VR_z U^\top UR_z^\top V^\top \\ &= VR_z R_z^\top V^\top \\ &= VV^\top \\ &= \mathbb{1} \end{aligned}$$

- ii.  $U$  and  $V$  are special orthogonal matrices with  $\det(U) = \det(V^\top) = 1$  (Slide 9, Chapter 5).

$$\det(R) = \det(UR_z^\top V^\top) = \underbrace{\det(U)}_1 \cdot \underbrace{\det(R_z^\top)}_1 \cdot \underbrace{\det(V^\top)}_1 = 1$$

2. (a)  $H = R + Tu^\top \Leftrightarrow R = H - Tu^\top$ .

$$\begin{aligned} E &= \hat{T}R \\ &= \hat{T}(H - Tu^\top) \\ &= \hat{T}H - \underbrace{\hat{T}T}_{=T \times T=0} u^\top \\ &= \hat{T}H \end{aligned}$$

(b)

$$\begin{aligned}
H^\top E + E^\top H &= H^\top (\hat{T}H) + (\hat{T}H)^\top H \\
&= H^\top (\hat{T}H) + H^\top \hat{T}^\top H \\
&= H^\top \hat{T}H - H^\top \hat{T}H \quad (\text{because } \hat{T} \text{ is skew-symmetric, i.e. } \hat{T}^\top = -\hat{T}) \\
&= 0
\end{aligned}$$

3. The notations below are as in Slide 6, Chapter 5. Note that the following slides deal with projected points in the normalized plane ( $Z = 1$ ), whereas here we assume pixel coordinates. The case of normalized coordinates is then just a special case with  $K = \mathbb{1}$ .

Rotation  $R$  and translation  $T$  are defined such that

$$g_{21} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

transforms a point from coordinate system 1 (CS1) to coordinate system 2 (CS2). This means that the inverse transformation (converting points from CS2 to CS1) is given by

$$g_{12} = g_{21}^{-1} = \begin{bmatrix} R^\top & -R^\top T \\ 0 & 1 \end{bmatrix}.$$

$o_1$  seen in CS1:  $[0 \ 0 \ 0 \ 1]^\top$  (homogeneous coordinates)

$o_1$  seen in CS2:  $g_{21} [0 \ 0 \ 0 \ 1]^\top = \begin{bmatrix} T \\ 1 \end{bmatrix}$

$e_2$  are the pixel coordinates of  $o_1$  projected into image 2:

$$\lambda_2 e_2 = K_2 \Pi_0 \begin{bmatrix} T \\ 1 \end{bmatrix} = K_2 T$$

$o_2$  seen in CS2:  $[0 \ 0 \ 0 \ 1]^\top$

$o_2$  seen in CS1:  $g_{12} [0 \ 0 \ 0 \ 1]^\top = \begin{bmatrix} -R^\top T \\ 1 \end{bmatrix}$

$e_1$  are the pixel coordinates of  $o_2$  projected into image 1:

$$\lambda_1 e_1 = K_1 \Pi_0 \begin{bmatrix} -R^\top T \\ 1 \end{bmatrix} = -K_1 R^\top T$$

$$\begin{aligned}
F e_1 &= \underbrace{(K_2^{-\top} \hat{T} R K_1^{-1})}_F \underbrace{\left(-\frac{1}{\lambda_1} K_1 R^\top T\right)}_{e_1} \\
&= -\frac{1}{\lambda_1} K_2^{-\top} \hat{T} R \underbrace{K_1^{-1} K_1}_\mathbb{1} R^\top T \\
&= -\frac{1}{\lambda_1} K_2^{-\top} \hat{T} \underbrace{R R^\top}_\mathbb{1} T \\
&= -\frac{1}{\lambda_1} K_2^{-\top} \underbrace{\hat{T} T}_{=T \times T = 0} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
e_2^\top F &= \underbrace{\left(\frac{1}{\lambda_2} K_2 T\right)^\top}_{e_2} \underbrace{\left(K_2^{-\top} \hat{T} R K_1^{-1}\right)}_F \\
&= \frac{1}{\lambda_2} T^\top \underbrace{K_2^\top K_2^{-\top}}_{\mathbb{1}} \hat{T} R K_1^{-1} \\
&= \frac{1}{\lambda_2} T^\top \hat{T} R K_1^{-1} \\
&= \frac{1}{\lambda_2} (\hat{T}^\top T)^\top R K_1^{-1} \\
&= \frac{1}{\lambda_2} (-\hat{T} T)^\top R K_1^{-1} \\
&= -\frac{1}{\lambda_2} (T \times T)^\top R K_1^{-1} \\
&= -\frac{1}{\lambda_2} 0 R K_1^{-1} \\
&= 0
\end{aligned}$$

## Part II: Practical Exercises

### Remarks

#### 1. Are there two or four possible solutions for $R$ and $T$ ?

The answer is four. You may find it confusing since in Slide 9 you are told to have two sets of  $R$  and  $T$ , while in Slide 14 it says there are four possible solutions. Recall that the essential matrix  $E$  is calculated by solving the equation  $\chi E^s = 0$  (at the bottom of Slide 11). Ideally the  $E^s$  you get should lie in the nullspace of  $\chi$ . However, since there are always errors in the point pairs you've chosen to make  $\chi$ , in practice it is very difficult to get an  $E^s$  that makes  $\chi E^s$  exactly 0. Instead we use the SVD of  $\chi$  to get the  $E^s$  which minimizes  $\|\chi E^s\|$ . In other words, the  $E^s$  we get will give us  $\chi E^s = \sigma$  with  $\sigma$  being some very small vector. Now think about this: what will happen if we turn the sign of the  $E^s$ ? We will get  $\chi(-E^s) = -\sigma$  which still gives us the smallest  $\|\chi E^s\|$  ( $\|\chi E^s\| = \|\sigma\| = \|-\sigma\| = \|\chi(-E^s)\|$ ).

Now we know that solving  $\chi E^s = 0$  for  $E^s$  will always give us two possible solutions  $E$  and  $-E$ . From each of them you can get two possible sets of  $R$  and  $T$  using the equations in Slide 9. Altogether we get four possible solutions. In practice we do the calculation in a little different way. We usually calculate according to Slide 14 to get four solutions out of  $E$ . The two extra solutions we get are nothing but the ones we should have got from  $-E$ .

#### 2. How to get the correct $R$ and $T$ from the possible solutions?

The criterion you should use to rule out the incorrect solutions is that all the reconstructed 3D points should have positive depth seen from **both of the camera coordinate systems**. In other words, both  $\lambda_1^j$  and  $\lambda_2^j$  in Slide 17 need to be positive.