▶Representation Theorem for Essential Matrices

Theorem

Let $\mathbf{E}$ *be a* 3×3 *matrix with SVD* $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$. *Then* $\mathbf{E}$ *is essential iff* $\mathbf{D} \simeq \lambda\, \mathrm{diag}(1, 1, 0)$.

$\lambda \neq 0$

Proof.

Direct:

If $\mathbf{E}$ is an essential matrix, then $\mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top$ in (12) must be orthogonal, hence $\mathbf{B} = \lambda\mathbf{I}$.

Converse:

$\mathbf{E}$ is fundamental with $\mathbf{D} = \lambda\, \mathrm{diag}(1, 1, 0)$ then we do not need $\mathbf{B}$ (as if $\mathbf{B} = \lambda\mathbf{I}$) and $\mathbf{U}(\mathbf{V}\mathbf{W})^\top$ is orthogonal, as required.

□

▶Essential Matrix Decomposition

We are decomposing $\mathbf{E}$ to $\mathbf{E} = [-\mathbf{t}_{21}]_\times \mathbf{R}_{21} = \mathbf{R}_{21}[-\mathbf{R}_{21}^\top \mathbf{t}]_\times$ [H&Z, sec. 9.6]

1. compute SVD of $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ and verify $\mathbf{D} = \lambda \operatorname{diag}(1, 1, 0)$
2. if $\det \mathbf{U} < 0$ transform it to $-\mathbf{U}$, do the same for $\mathbf{V}$ — the overall sign is dropped
3. compute

$$\mathbf{R}_{21} = \mathbf{U} \underbrace{\begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{W}} \mathbf{V}^\top, \quad \mathbf{t}_{21} = -\beta\, \mathbf{u}_3, \qquad |\alpha| = 1, \quad \beta \neq 0 \tag{13}$$

Notes

- $\mathbf{U}(\mathbf{V}\mathbf{W})^\top \mathbf{v}_3 = \cdots = \mathbf{u}_3$
- $\mathbf{t}_{21}$ is recoverable up to scale β and direction $\operatorname{sign} \beta$
- the result for $\mathbf{R}_{21}$ is unique up to $\alpha = \pm 1$ — despite non-uniqueness of SVD
- change of sign in $\mathbf{W}$ rotates the solution by 180° about $\mathbf{t}$

  $\mathbf{R}_1 = \mathbf{U}\mathbf{W}\mathbf{V}^\top$, $\mathbf{R}_2 = \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top \Rightarrow \mathbf{T} = \mathbf{R}_2\mathbf{R}_1^\top = \cdots = \mathbf{U} \operatorname{diag}(-1, -1, 1)\mathbf{U}^\top$ which is a rotation by 180° about $\mathbf{u}_3 = \mathbf{t}_{21}$:

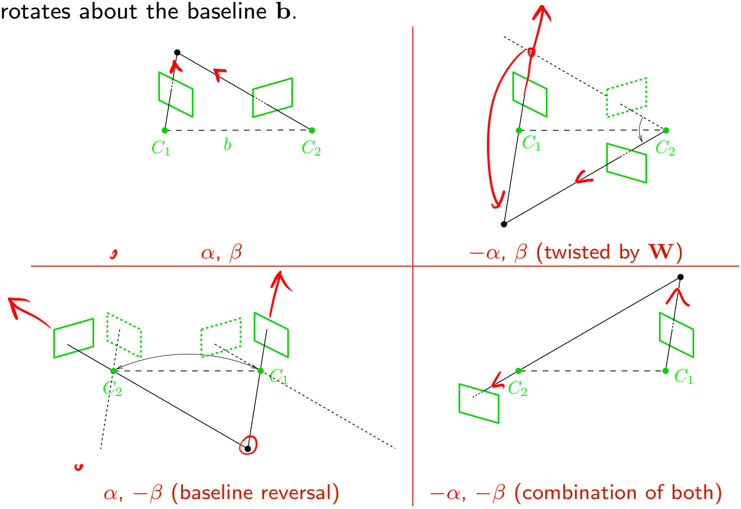
$$U^T = \begin{bmatrix} u_1^T \\ u_2^T \\ u_3^T \end{bmatrix} u_3$$

$$\mathbf{U} \operatorname{diag}(-1, -1, 1)\mathbf{U}^\top \mathbf{u}_3 = \mathbf{U} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{u}_3$$

- 4 solution sets for 4 sign combinations of α, β — see next for geometric interpretation

▶Four Solutions to Essential Matrix Decomposition

Transform the world coordinate system so that the origin is in Camera 2. Then $\mathbf{t}_{21} = -\mathbf{b}$ and $\mathbf{W}$ rotates about the baseline $\mathbf{b}$. →73



- chirality constraint: all 3D points are in front of both cameras
- this singles-out the upper left case [H&Z, Sec. 9.6.3]

►7-Point Algorithm for Estimating Fundamental Matrix

Problem: Given a set $\{(x_i, y_i)\}_{i=1}^{k}$ of $k = 7$ correspondences, estimate f. m. $\mathbf{F}$.

$$\underline{\mathbf{y}}_i^\top \mathbf{F}\, \underline{\mathbf{x}}_i = 0, \quad i = 1, \dots, k, \qquad \underline{\text{known}}: \quad \underline{\mathbf{x}}_i = (u_i^1, v_i^1, 1), \quad \underline{\mathbf{y}}_i = (u_i^2, v_i^2, 1)$$

terminology: correspondence = truth, later: match = algorithm's result; hypothesized corresp.

Solution:

$$\mathbf{D} = \begin{bmatrix} u_1^1 u_1^2 & u_1^1 v_1^2 & u_1^1 & u_1^2 v_1^1 & v_1^1 v_1^2 & v_1^1 & u_1^2 & v_1^2 & 1 \\ u_2^1 u_2^2 & u_2^1 v_2^2 & u_2^1 & u_2^2 v_2^1 & v_2^1 v_2^2 & v_2^1 & u_2^2 & v_2^2 & 1 \\ u_3^1 u_3^2 & u_3^1 v_3^2 & u_3^1 & u_3^2 v_3^1 & v_3^1 v_3^2 & v_3^1 & u_3^2 & v_3^2 & 1 \\ \vdots & & & & & & & & \vdots \\ u_k^1 u_k^2 & u_k^1 v_k^2 & u_k^1 & u_k^2 v_k^1 & v_k^1 v_k^2 & v_k^1 & u_k^2 & v_k^2 & 1 \end{bmatrix} \qquad \mathbf{D} \in \mathbb{R}^{k,9}$$

$$\mathbf{D}\, \mathrm{vec}(\mathbf{F}) = \mathbf{0}, \qquad \mathrm{vec}(\mathbf{F}) = \begin{bmatrix} f_{11} & f_{21} & f_{31} & \dots & f_{33} \end{bmatrix}^\top, \qquad \mathrm{vec}(\mathbf{F}) \in \mathbb{R}^9,$$

- for $k = 7$ we have a rank-deficient system, the null-space of $\mathbf{D}$ is 2-dimensional
- but we know that $\det \mathbf{F} = 0$, hence
  1. find a basis of the null space of $\mathbf{D}$: $\mathbf{F}_1$, $\mathbf{F}_2$ — by SVD or QR factorization
  2. get up to 3 real solutions for α_i from

     $$\det(\alpha \mathbf{F}_1 + (1 - \alpha)\mathbf{F}_2) = 0 \qquad \text{cubic equation in } \alpha$$

  3. get up to 3 fundamental matrices $\mathbf{F} = \alpha_i \mathbf{F}_1 + (1 - \alpha_i)\mathbf{F}_2$ — (check $\mathrm{rank}\, \mathbf{F} = 2$)

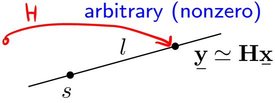
- the result may depend on image transformations
- normalization improves conditioning →87
- this gives a good starting point for the full algorithm →106
- dealing with mismatches need not be a part of the 7-point algorithm →107

▶Degenerate Configurations for Fundamental Matrix Estimation

When is $\mathbf{F}$ not uniquely determined from any number of correspondences? [H&Z, Sec. 11.9]

1. when images are related by homography
   a) camera centers coincide $C_1 = C_2$: $\quad \mathbf{H} = \mathbf{K}_2 \mathbf{R}_{21} \mathbf{K}_1^{-1}$
   b) camera moves but all 3D points lie in a plane $(\mathbf{n}, d)$: $\quad \mathbf{H} = \mathbf{K}_2 (\mathbf{R}_{21} - \mathbf{t}_{21} \mathbf{n}^\top / d) \mathbf{K}_1^{-1}$

   - in both cases: epipolar geometry is not defined
   - we do get an $\mathbf{F}$ from the 7-point algorithm but it is of the form of $\mathbf{F} = [\underline{\mathbf{s}}]_\times \mathbf{H}$ with $\underline{\mathbf{s}}$ arbitrary (nonzero) note that $[\underline{\mathbf{s}}]_\times \mathbf{H} \simeq \mathbf{H}' [\underline{\mathbf{s}}']_\times$ →72



   - correspondence $x \leftrightarrow y$
   - y is the image of x: $\underline{\mathbf{y}} \simeq \mathbf{H}\underline{\mathbf{x}}$
   - a necessary condition: $y \in l$, $\quad \underline{\mathbf{l}} \simeq \underline{\mathbf{s}} \times \mathbf{H}\underline{\mathbf{x}}$ arbitrary $\underline{\mathbf{s}}$

$$0 = \underline{\mathbf{y}}^\top (\underline{\mathbf{s}} \times \mathbf{H}\underline{\mathbf{x}}) = \underline{\mathbf{y}}^\top [\underline{\mathbf{s}}]_\times \mathbf{H}\underline{\mathbf{x}}$$

2. both camera centers and all 3D points lie on a ruled quadric

   hyperboloid of one sheet, cones, cylinders, two planes

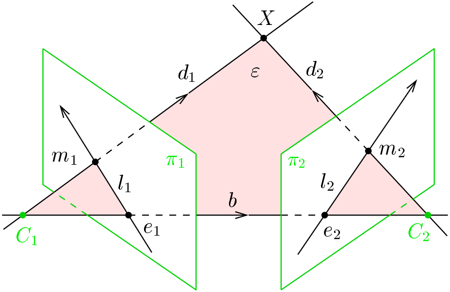
   - there are 3 solutions for $\mathbf{F}$

notes

- estimation of $\mathbf{E}$ can deal with planes: $[\underline{\mathbf{s}}]_\times \mathbf{H} = [\underline{\mathbf{s}}]_\times (\mathbf{R}_{21} - \mathbf{t}_{21} \mathbf{n}^\top / d)$ has equal ~~eigenvalues~~ singular values iff $\underline{\mathbf{s}} = \mathbf{t}_{21}$, the decomposition works (nonunique, as before) ⊛ P1; 1pt for a proof
- a complete treatment with additional degenerate configurations in [H&Z, sec. 22.2]
- a stronger epipolar constraint could reject some configurations

A Note on Oriented Epipolar Constraint

- a tighter epipolar constraint preserves orientations
- requires all points and cameras be on the same side of the plane at infinity



$$\underline{\mathbf{e}}_2 \times \underline{\mathbf{m}}_2 \overset{+}{\sim} \mathbf{F}\,\underline{\mathbf{m}}_1$$

notation: $\underline{\mathbf{m}} \overset{+}{\sim} \underline{\mathbf{n}}$ means $\underline{\mathbf{m}} = \lambda \underline{\mathbf{n}}$, $\lambda > 0$

- note that the constraint is not invariant to the change of either sign of $\underline{\mathbf{m}}_i$
- all 7 correspondence in 7-point alg. must have the same sign — see later
- this may help reject some wrong matches, see →107 — [Chum et al. 2004]
- an even more tight constraint: scene points in front of both cameras — expensive

  this is called chirality constraint

▶5-Point Algorithm for Relative Camera Orientation

Problem: Given $\{m_i, m_i'\}_{i=1}^{5}$ corresponding image points and calibration matrix $\mathbf{K}$, recover the camera motion $\mathbf{R}$, $\mathbf{t}$.

Obs:

1. $\mathbf{E}$ – 8 numbers
2. $\mathbf{R}$ – 3DOF, $\mathbf{t}$ – we can recover 2DOF only, in total 5 DOF $\to$ we need 3 constraints on $\mathbf{E}$
3. $\mathbf{E}$ essential iff it has two equal singular values and the third is zero

This gives an equation system:

$$\underline{\mathbf{v}}_i^\top \mathbf{E}\, \underline{\mathbf{v}}_i' = 0 \qquad \text{5 linear constraints } (\underline{\mathbf{v}} \simeq \mathbf{K}^{-1}\underline{\mathbf{m}})$$

$$\det \mathbf{E} = 0 \qquad \text{1 cubic constraint}$$

$$\mathbf{E}\mathbf{E}^\top\mathbf{E} - \frac{1}{2}\operatorname{tr}(\mathbf{E}\mathbf{E}^\top)\mathbf{E} = \mathbf{0} \qquad \text{9 cubic constraints, 2 independent}$$

⊛ P1; 1pt: verify this equation from $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$, $\mathbf{D} = \lambda\,\mathrm{diag}(1, 1, 0)$

1. estimate $\mathbf{E}$ by SVD from $\underline{\mathbf{v}}_i^\top \mathbf{E}\, \underline{\mathbf{v}}_i' = 0$ by the null-space method,
2. this gives $\mathbf{E} = x\mathbf{E}_1 + y\mathbf{E}_2 + z\mathbf{E}_3 + \mathbf{E}_4$ red
3. at most 10 (complex) solutions for x, y, z from the cubic constraints

- when all 3D points lie on a plane: at most 2 solutions (twisted-pair) can be disambiguated in 3 views or by chirality constraint (→79) unless all 3D points are closer to one camera
- 6-point problem for unknown f [Kukelova et al. BMVC 2008]
- resources at http://cmp.felk.cvut.cz/minimal/5_pt_relative.php

▶The Triangulation Problem

Problem: Given cameras $\mathbf{P}_1$, $\mathbf{P}_2$ and a correspondence $x \leftrightarrow y$ compute a 3D point $\mathbf{X}$ projecting to x and y

$$\lambda_1\, \underline{\mathbf{x}} = \mathbf{P}_1 \underline{\mathbf{X}}, \qquad \lambda_2\, \underline{\mathbf{y}} = \mathbf{P}_2 \underline{\mathbf{X}}, \qquad \underline{\mathbf{x}} = \begin{bmatrix} u^1 \\ v^1 \\ 1 \end{bmatrix}, \qquad \underline{\mathbf{y}} = \begin{bmatrix} u^2 \\ v^2 \\ 1 \end{bmatrix}, \qquad \mathbf{P}_i = \begin{bmatrix} (\mathbf{p}_1^i)^\top \\ (\mathbf{p}_2^i)^\top \\ (\mathbf{p}_3^i)^\top \end{bmatrix}$$

Linear triangulation method

$$u^1\,(\mathbf{p}_3^1)^\top \underline{\mathbf{X}} = (\mathbf{p}_1^1)^\top \underline{\mathbf{X}}, \qquad u^2\,(\mathbf{p}_3^2)^\top \underline{\mathbf{X}} = (\mathbf{p}_1^2)^\top \underline{\mathbf{X}},$$

$$v^1\,(\mathbf{p}_3^1)^\top \underline{\mathbf{X}} = (\mathbf{p}_2^1)^\top \underline{\mathbf{X}}, \qquad v^2\,(\mathbf{p}_3^2)^\top \underline{\mathbf{X}} = (\mathbf{p}_2^2)^\top \underline{\mathbf{X}},$$

Gives

$$\mathbf{D}\underline{\mathbf{X}} = \mathbf{0}, \qquad \mathbf{D} = \begin{bmatrix} u^1\,(\mathbf{p}_3^1)^\top - (\mathbf{p}_1^1)^\top \\ v^1\,(\mathbf{p}_3^1)^\top - (\mathbf{p}_2^1)^\top \\ u^2\,(\mathbf{p}_3^2)^\top - (\mathbf{p}_1^2)^\top \\ v^2\,(\mathbf{p}_3^2)^\top - (\mathbf{p}_2^2)^\top \end{bmatrix}, \qquad \mathbf{D} \in \mathbb{R}^{4,4}, \quad \underline{\mathbf{X}} \in \mathbb{R}^4 \tag{14}$$

$\min \| D \underline{X} \|^2$

- back-projected rays will generally not intersect due to image error, see next
- using Jack-knife (→63) not recommended — sensitive to small error
- we will use SVD (→85)
- but the result will not be invariant to projective frame
  replacing $\mathbf{P}_1 \mapsto \mathbf{P}_1\mathbf{H}$, $\mathbf{P}_2 \mapsto \mathbf{P}_2\mathbf{H}$ does not always result in $\underline{\mathbf{X}} \mapsto \mathbf{H}^{-1}\underline{\mathbf{X}}$
- note the homogeneous form in (14) can represent points at infinity

▶The Least-Squares Triangulation by SVD

- if $\mathbf{D}$ is full-rank we may minimize the algebraic least-squares error

$$\boldsymbol{\varepsilon}^2(\underline{\mathbf{X}}) = \|\mathbf{D}\underline{\mathbf{X}}\|^2 \quad \text{s.t.} \quad \|\underline{\mathbf{X}}\| = 1, \qquad \underline{\mathbf{X}} \in \mathbb{R}^4$$

- let $\mathbf{D}_i$ be the i-th row of $\mathbf{D}$, then

$$\|\mathbf{D}\underline{\mathbf{X}}\|^2 = \sum_{i=1}^{4} \underbrace{(\mathbf{D}_i\, \underline{\mathbf{X}})^2}_{a^\top \cdot a} = \sum_{i=1}^{4} \underline{\mathbf{X}}^\top \underbrace{\mathbf{D}_i^\top \mathbf{D}_i}_{4\times 4}\, \underline{\mathbf{X}} = \underline{\mathbf{X}}^\top \mathbf{Q}\, \underline{\mathbf{X}}, \text{ where } \underline{\mathbf{Q}} = \sum_{i=1}^{4} \mathbf{D}_i^\top \mathbf{D}_i = \mathbf{D}^\top \mathbf{D} \in \mathbb{R}^{4,4}$$

- we write the SVD of $\mathbf{Q}$ as $\mathbf{Q} = \sum_{j=1}^{4} \sigma_j^2\, \mathbf{u}_j \mathbf{u}_j^\top$, in which [Golub & van Loan 2013, Sec. 2.5]

$Q = UDU^\top$

$$\sigma_1^2 \geq \cdots \geq \sigma_4^2 \geq 0 \quad \text{and} \quad \mathbf{u}_l^\top \mathbf{u}_m = \begin{cases} 0 & \text{if } l \neq m \\ 1 & \text{otherwise} \end{cases}$$

- then $\underline{\mathbf{X}} = \arg \min_{\mathbf{q}, \|\mathbf{q}\|=1} \mathbf{q}^\top \mathbf{Q}\, \mathbf{q} = \mathbf{u}_4$

Proof (by contradiction).

$$\mathbf{q}^\top \mathbf{Q}\, \mathbf{q} = \sum_{j=1}^{4} \sigma_j^2 (\mathbf{q}^\top \mathbf{u}_j)\, \mathbf{u}_j^\top \mathbf{q} = \sum_{j=1}^{4} \sigma_j^2\, (\mathbf{u}_j^\top \mathbf{q})^2$$ is a sum of non-negative elements $0 \leq (\mathbf{u}_j^\top \mathbf{q})^2 \leq 1$

$\|q\| = 1$ and $\cos^2\alpha \cdot \sigma_4^2 + \sin^2\alpha \underbrace{\sum_{j=1}^{3} \sigma_j^2 (u_j^\top \bar{q})^2}_{\geq \sigma_4^2} \geq \sigma_4^2$

Let $\mathbf{q} = \cancel{\mathbf{u}_4 + \bar{\mathbf{q}}}$ s.t. $\left\{ \begin{array}{l} \bar{\mathbf{q}} \perp \mathbf{u}_4 \\ \|\bar{q}\| = 1 \end{array} \right.$, then

$q = \cos\alpha\, u_4 + \sin\alpha\, \bar{q}$

$0 \leq \alpha \leq 1$

$$\mathbf{q}^\top \mathbf{Q}\, \mathbf{q} \overset{\dots =}{=} \sigma_4^2 + \sum_{j=1}^{3} \sigma_j^2\, (\mathbf{u}_j^\top \bar{\mathbf{q}})^2 \geq \sigma_4^2$$

- if $\sigma_4 \ll \sigma_3$, there is a unique solution $\underline{\mathbf{X}} = \mathbf{u}_4$ with residual error $(\mathbf{D}\,\underline{\mathbf{X}})^2 = \sigma_4^2$
  the quality (conditioning) of the solution may be expressed as $q = \sigma_3/\sigma_4$ (greater is better)

Matlab code for the least-squares solver:

```
[U,O,V] = svd(D);
X = V(:,end);
q = O(3,3)/O(4,4);
```

⊛ P1; 1pt: Why did we decompose $\mathbf{D}$ and not $\mathbf{Q} = \mathbf{D}^\top\mathbf{D}$?