3D Computer Vision

Radim Šára   Martin Matoušek

Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

https://cw.fel.cvut.cz/wiki/courses/tdv/start

http://cmp.felk.cvut.cz
mailto:sara@cmp.felk.cvut.cz
phone ext. 7203

rev. October 25, 2022

Open Informatics Master's Course

►Representation Theorem for Fundamental Matrices

Def: $\mathbf{F}$ is fundamental when $\mathbf{F} \simeq \mathbf{H}^{-\top}[\underline{\mathbf{e}}_1]_\times$, where $\mathbf{H}$ is regular and $\underline{\mathbf{e}}_1 \simeq \operatorname{null}\mathbf{F} \neq \mathbf{0}$.

Theorem: A 3×3 matrix $\mathbf{A}$ is fundamental iff it is of rank 2.

Proof.

Direct: By the geometry, $\mathbf{H}$ is full-rank, $\underline{\mathbf{e}}_1 \neq \mathbf{0}$, hence $\mathbf{H}^{-\top}[\underline{\mathbf{e}}_1]_\times$ is a 3×3 matrix of rank 2.

Converse:

1. let $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ be the SVD of $\mathbf{A}$ of rank 2; then $\mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2, 0)$, $\lambda_1 \geq \lambda_2 > 0$
2. we write $\mathbf{D} = \mathbf{B}\mathbf{C}$, where $\mathbf{B} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, $\mathbf{C} = \operatorname{diag}(1, 1, 0)$, $\lambda_3 > 0$
3. then $\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{C}\mathbf{V}^\top = \mathbf{U}\mathbf{B}\mathbf{C}\underbrace{\mathbf{W}\mathbf{W}^\top}_{\mathbf{I}}\mathbf{V}^\top$ with $\mathbf{W}$ rotation matrix
4. we look for a rotation mtx $\mathbf{W}$ that maps $\mathbf{C}$ to a skew-symmetric $\mathbf{S}$, i.e. $\mathbf{S} = \mathbf{C}\mathbf{W}$, if any
5. then $\mathbf{W} = \begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $|\alpha| = 1$, and $\mathbf{S} = \mathbf{C}\mathbf{W} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \cdots = [\mathbf{s}]_\times$, where $\mathbf{s} = (0, 0, 1)$
6. we write

$\mathbf{v}_3$ – 3rd column of $\mathbf{V}$, $\mathbf{u}_3$ – 3rd column of $\mathbf{U}$

$$\mathbf{A} = \mathbf{U}\mathbf{B}[\mathbf{s}]_\times \mathbf{W}^\top \mathbf{V}^\top \overset{\circledast 1}{=} \cdots = \underbrace{\mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top}_{\simeq \mathbf{H}^{-\top}} [\mathbf{v}_3]_\times \overset{\rightarrow 76/9}{\simeq} \underbrace{[\mathbf{H}\mathbf{v}_3]_\times}_{\simeq [\mathbf{u}_3]_\times} \mathbf{H}, \qquad (12)$$

7. $\mathbf{H}$ regular, $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$, $\mathbf{u}_3\mathbf{A} = \mathbf{0}$ for $\mathbf{v}_3 \neq \mathbf{0}$, $\mathbf{u}_3 \neq \mathbf{0}$ □

- we also got a (non-unique: α, λ_3) decomposition formula for fundamental matrices
- it follows there is no constraint on $\mathbf{F}$ except for the rank

▶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* $\mathbf{D} \simeq \mathrm{diag}(1, 1, 0)$.

Proof.

Direct:

If $\mathbf{E}$ is an essential matrix, then the epipolar homography matrix is a rotation matrix (→78), hence $\mathbf{H}^{-\top} \simeq \mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top$ in (12) must be (1) diagonal, and (2) (λ-scaled) ~~orthogonal~~. normal

It follows $\mathbf{B} = \lambda \mathbf{I}$. note this fixed the missing λ_3 in (12)

Then

$$\mathbf{R}_{21} = \mathbf{H}^{-\top} \simeq \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top \simeq \mathbf{U}\mathbf{W}\mathbf{V}^\top$$

Converse:

$\mathbf{E}$ is fundamental with

$$\mathbf{D} = \mathrm{diag}(\lambda, \lambda, 0) = \underbrace{\lambda\mathbf{I}}_{\mathbf{B}}\ \underbrace{\mathrm{diag}(1, 1, 0)}_{\mathbf{D}}$$

then $\mathbf{B} = \lambda \mathbf{I}$ in (12) and $\mathbf{U}(\mathbf{V}\mathbf{W})^\top$ is orthogonal, as required. □

►Essential Matrix Decomposition

We are decomposing $\mathbf{E}$ to $\mathbf{E} \simeq [\mathbf{u}_3]_\times \mathbf{H} \simeq [-\mathbf{t}_{21}]_\times \mathbf{R}_{21} = \mathbf{R}_{21}[-\mathbf{R}_{21}^\top \mathbf{t}_{21}]_\times \simeq \mathbf{H}^{-\top}[\mathbf{v}_3]_\times$ [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. ensure $\mathbf{U}$, $\mathbf{V}$ are rotation matrices by $\mathbf{U} \mapsto \det(\mathbf{U})\mathbf{U}$, $\mathbf{V} \mapsto \det(\mathbf{V})\mathbf{V}$
3. compute

$$\mathbf{R}_{21} = \mathbf{U} \underbrace{\begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{W}} \mathbf{V}^\top, \quad \mathbf{t}_{21} \overset{(12)}{=} -\beta\, \mathbf{u}_3, \qquad |\alpha| = 1, \quad \beta \neq 0 \tag{13}$$

Notes

- $\mathbf{v}_3 \simeq \mathbf{R}_{21}^\top \mathbf{t}_{21}$ by (12), hence $\mathbf{R}_{21}\mathbf{v}_3 \simeq \mathbf{t}_{21} \simeq \mathbf{u}_3$ since it must fall in left null space by $\mathbf{E} \simeq [\mathbf{u}_3]_\times \mathbf{R}_{21}$
- $\mathbf{t}_{21}$ 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
- the change of sign in α rotates the solution by $180°$ about $\mathbf{t}_{21}$

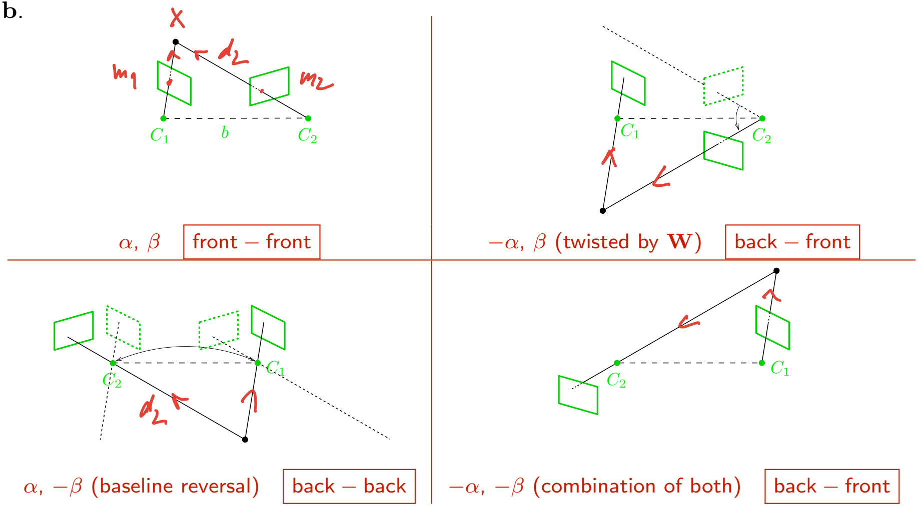
  $\mathbf{R}(\alpha) = \mathbf{U}\mathbf{W}\mathbf{V}^\top \;\Rightarrow\; \mathbf{R}(-\alpha) = \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top \Rightarrow \mathbf{T} = \mathbf{R}(-\alpha)\mathbf{R}^\top(\alpha) = \cdots = \mathbf{U}\operatorname{diag}(-1, -1, 1)\mathbf{U}^\top$
  which is a rotation by $180°$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

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

- 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 $\mathbf{b}$. →77



How to disambiguate?

- use the chirality constraint: all 3D points are in front of both cameras
- this singles-out the upper left case: front-front

[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 $k = 7$ finite correspondences, estimate f. m. $\mathbf{F}$.

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

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

Solution:

$$\underline{\mathbf{y}}_i^\top \mathbf{F}\, \underline{\mathbf{x}}_i = (\underline{\mathbf{y}}_i \underline{\mathbf{x}}_i^\top) : \mathbf{F} = \left(\operatorname{vec}(\underline{\mathbf{y}}_i \underline{\mathbf{x}}_i^\top)\right)^\top \operatorname{vec}(\mathbf{F}),$$

rotation property of matrix trace →71

$$\operatorname{vec}(\mathbf{F}) = \begin{bmatrix} f_{11} & f_{21} & f_{31} & \ldots & f_{33} \end{bmatrix}^\top \in \mathbb{R}^9$$

column vector from matrix

$$\mathbf{D} = \begin{bmatrix} \left(\operatorname{vec}(\underline{\mathbf{y}}_1 \underline{\mathbf{x}}_1^\top)\right)^\top \\ \left(\operatorname{vec}(\underline{\mathbf{y}}_2 \underline{\mathbf{x}}_2^\top)\right)^\top \\ \left(\operatorname{vec}(\underline{\mathbf{y}}_3 \underline{\mathbf{x}}_3^\top)\right)^\top \\ \vdots \\ \left(\operatorname{vec}(\underline{\mathbf{y}}_k \underline{\mathbf{x}}_k^\top)\right)^\top \end{bmatrix} = \begin{bmatrix} u_1^1 u_1^2 & u_1^1 v_1^2 & u_1^1 & u_1^2 v_1^1 & v_1^1 v_1^2 & v_1^1 & u_1^2 & v_1^2 & 1 \\ u_2^1 u_2^2 & u_2^1 v_2^2 & u_2^1 & u_2^2 v_2^1 & v_2^1 v_2^2 & v_2^1 & u_2^2 & v_2^2 & 1 \\ u_3^1 u_3^2 & u_3^1 v_3^2 & u_3^1 & u_3^2 v_3^1 & v_3^1 v_3^2 & v_3^1 & u_3^2 & v_3^2 & 1 \\ \vdots & & & & & & & & \vdots \\ u_k^1 u_k^2 & u_k^1 v_k^2 & u_k^1 & u_k^2 v_k^1 & v_k^1 v_k^2 & v_k^1 & u_k^2 & v_k^2 & 1 \end{bmatrix} \in \mathbb{R}^{k,9}$$

$$\mathbf{D}\operatorname{vec}(\mathbf{F}) = \mathbf{0}$$

▶7-Point Algorithm Continued

$$\mathbf{D}\,\mathrm{vec}(\mathbf{F}) = \mathbf{0}, \quad \mathbf{D} \in \mathbb{R}^{k,9}$$

- 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 from
     $$\det(\alpha \mathbf{F}_1 + (1-\alpha)\mathbf{F}_2) = 0$$
     cubic equation in α
  3. get up to 3 fundamental matrices $\mathbf{F}_i = \alpha_i \mathbf{F}_1 + (1-\alpha_i)\mathbf{F}_2$
  4. if $\mathrm{rank}\,\mathbf{F}_i < 2$ for all $i = 1,2,3$ then fail

- the result may depend on image (domain) transformations
- normalization improves conditioning →92
- this gives a good starting point for the full algorithm →111
- dealing with mismatches need not be a part of the 7-point algorithm →112

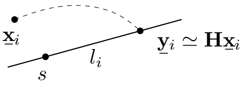
▶Degenerate Configurations for Fundamental Matrix Estimation

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 $\mathbf{t}_{21} = 0$: $\mathbf{H} = \mathbf{K}_2\mathbf{R}_{21}\mathbf{K}_1^{-1}$ — $\mathbf{H}$ – as in epipolar homography
   b) camera moves but all 3D points lie in a plane $(\mathbf{n}, d)$: $\mathbf{H} = \mathbf{K}_2(\mathbf{R}_{21} - \mathbf{t}_{21}\mathbf{n}^\top/d)\mathbf{K}_1^{-1}$

- in either case: epipolar geometry is not defined
- we get an arbitrary solution from the 7-point algorithm, in the form of $\mathbf{F} = [\underline{\mathbf{s}}]_\times \mathbf{H}$

note that $[\underline{\mathbf{s}}]_\times \mathbf{H} \simeq \mathbf{H}'[\underline{\mathbf{s}}']_\times$ →76



- given (arbitrary, fixed) point $\underline{\mathbf{s}}$
- and correspondence $x_i \leftrightarrow y_i$
- y_i is the image of x_i: $\underline{\mathbf{y}}_i \simeq \mathbf{H}\underline{\mathbf{x}}_i$
- a necessary condition: $y_i \in l_i$, $\quad \underline{\mathbf{l}}_i \simeq \underline{\mathbf{s}} \times \mathbf{H}\underline{\mathbf{x}}_i$

$$0 = \underline{\mathbf{y}}_i^\top(\underline{\mathbf{s}} \times \mathbf{H}\underline{\mathbf{x}}_i) = \underline{\mathbf{y}}_i^\top[\underline{\mathbf{s}}]_\times \mathbf{H}\underline{\mathbf{x}}_i \quad \text{for any } \underline{\mathbf{x}}_i, \underline{\mathbf{y}}_i, \underline{\mathbf{s}} \text{ (!)}$$

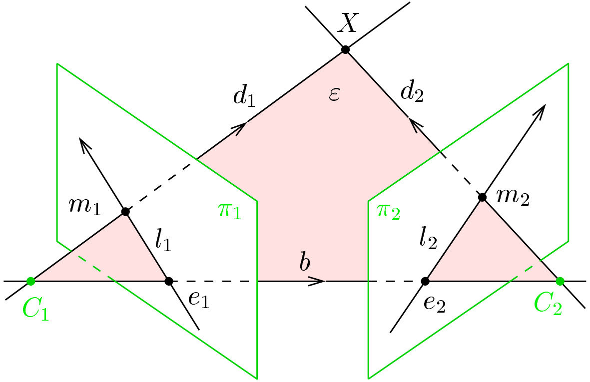
2. both camera centers and all 3D points lie on a ruled quadric
   hyperboloid of one sheet, cones, cylinders, two planes
   - there are 3 solutions for $\mathbf{F}$

notes

- estimation of $\mathbf{E}$ can deal with planes: $[\underline{\mathbf{s}}]_\times \mathbf{H}$ is essential, then $\mathbf{H} = \mathbf{R} - \mathbf{t}\mathbf{n}^\top/d$, and $\underline{\mathbf{s}} \simeq \mathbf{t}$ not arbitrary
- a complete treatment with additional degenerate configurations in [H&Z, sec. 22.2]
- a stronger epipolar constraint could reject some configurations (see next)

A Note on Oriented Epipolar Constraint

- a tighter epipolar constraint that 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 $\underline{\mathbf{m}} = \lambda \underline{\mathbf{n}}$, $\lambda > 0$

- we can read the constraint as $(\underline{\mathbf{e}}_2 \times \underline{\mathbf{m}}_2) \overset{+}{\sim} \mathbf{H}_e^{-\top} (\underline{\mathbf{e}}_1 \times \underline{\mathbf{m}}_1)$
- 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 →112 — [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. $\mathbf{E}$ – homogeneous 3×3 matrix; 9 numbers up to scale
2. $\mathbf{R}$ – 3 DOF, $\mathbf{t}$ – 2 DOF only, in total 5 DOF $\rightarrow$ we need $9 - 1 - 5 = 3$ constraints on $\mathbf{E}$
3. idea: $\mathbf{E}$ essential iff it has two equal singular values and the third is zero →81

This gives an equation system:

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

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

$\lambda_1 = \lambda_2, \lambda_3 = 0 \Leftrightarrow$

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

⊛ P1; 1pt: verify the last 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 — 4D null space
2. this gives $\mathbf{E} \simeq x\mathbf{E}_1 + y\mathbf{E}_2 + z\mathbf{E}_3 + \mathbf{E}_4$
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 real solutions (twisted-pair) — can be disambiguated in 3 views or by chirality constraint (→83) unless all 3D points are closer to one camera
- 6-point problem for unknown f — [Kukelova et al. BMVC 2008]
- resources at `http://aag.ciirc.cvut.cz/minimal/`

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 after eliminating λ_1, λ_2

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

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

Gives

$\|D\underline{X}\| \to$ small

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

- typically, $\mathbf{D}$ has full rank (!)
- what else: back-projected rays will generally not intersect due to image error, see next
- what else: using Jack-knife ($\rightarrow$63) not recommended — sensitive to small error
- idea: we will step back and use SVD ($\rightarrow$90)
- 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 $\underline{\mathbf{X}}$ at infinity

- 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 $\mathbf{D}$ taken as a column vector, then

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

- 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]

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

- then $\underline{\mathbf{X}} = \arg \min_{\mathbf{q}, \|\mathbf{q}\|=1} \mathbf{q}^\top \mathbf{Q}\, \mathbf{q} = \mathbf{u}_4$ — the last column of the $\mathbf{U}$ matrix from $\mathrm{SVD}(\mathbf{D}^\top \mathbf{D})$

min $q^\top Q q = \sigma_4^2$

Proof (by contradiction).

Let $\bar{\mathbf{q}} = \sum_{i=1}^{4} a_i \mathbf{u}_i$ s.t. $\sum_{i=1}^{4} a_i^2 = 1$, then $\|\bar{\mathbf{q}}\| = 1$, as desired, and

$$\bar{\mathbf{q}}^\top \mathbf{Q}\, \bar{\mathbf{q}} = \sum_{j=1}^{4} \sigma_j^2\, \bar{\mathbf{q}}^\top \mathbf{u}_j \mathbf{u}_j^\top \bar{\mathbf{q}} = \sum_{j=1}^{4} \sigma_j^2\, (\mathbf{u}_j^\top \bar{\mathbf{q}})^2 = \cdots = \sum_{j=1}^{4} a_j^2 \sigma_j^2 \geq \sum_{j=1}^{4} a_j^2 \sigma_4^2 = \left(\sum_{j=1}^{4} a_j^2\right) \sigma_4^2 = \sigma_4^2$$

$\Rightarrow \bar{q} = u_4$

since $\sigma_j \geq \sigma_4$ □

▶cont'd

- if $\sigma_4 \ll \sigma_3$, there is a unique solution $\underline{\mathbf{X}} = \mathbf{u}_4$ with residual error $(\mathbf{D}\,\underline{\mathbf{X}})^2 = \sigma_4^2$

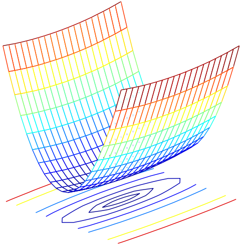
  the quality (conditioning) of the solution may be expressed as $q = \sigma_3/\sigma_4$ (greater is better)

Matlab code for the least-squares solver:

```
[U,O,V] = svd(D);
X = V(:,end);
q = sqrt(O(end-1,end-1)/O(end,end));
```

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

▶Numerical Conditioning



- The equation $\mathbf{D}\underline{\mathbf{X}} = \mathbf{0}$ in (14) may be ill-conditioned for numerical computation, which results in a poor estimate for $\underline{\mathbf{X}}$.

Why: on a row of $\mathbf{D}$ there are big entries together with small entries, e.g. of orders projection centers in mm, image points in px

$$\begin{bmatrix} 10^3 & 0 & 10^3 & 10^6 \\ 0 & 10^3 & 10^3 & 10^6 \\ 10^3 & 0 & 10^3 & 10^6 \\ 0 & 10^3 & 10^3 & 10^6 \end{bmatrix}$$

Quick fix:

1. re-scale the problem by a regular diagonal conditioning matrix $\mathbf{S} \in \mathbb{R}^{4,4}$

$$\mathbf{0} = \mathbf{D}\,\underline{\mathbf{X}} = \mathbf{D}\,\mathbf{S}\,\mathbf{S}^{-1}\underline{\mathbf{X}} = \bar{\mathbf{D}}\,\underline{\bar{\mathbf{X}}}$$

   choose $\mathbf{S}$ to make the entries in $\hat{\mathbf{D}}$ all smaller than unity in absolute value:

$$\mathbf{S} = \operatorname{diag}(10^{-3}, 10^{-3}, 10^{-3}, 10^{-6})$$

```
S = diag(1./max(abs(D), [], 1))
```

2. solve for $\underline{\bar{\mathbf{X}}}$ as before
3. get the final solution as $\underline{\mathbf{X}} = \mathbf{S}\,\underline{\bar{\mathbf{X}}}$

- when SVD is used in camera resection, conditioning is essential for success →62

Thank You