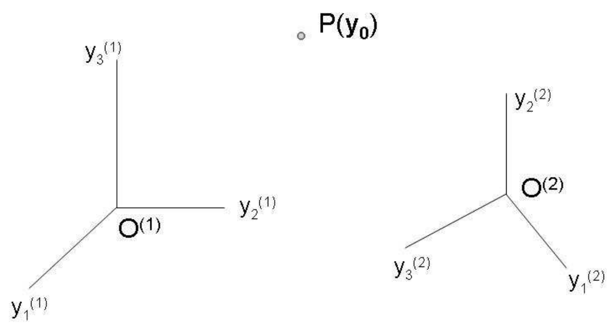
Similarity transformation in 3D between two matched points patterns.



coordinate system 1 coordinate system 2

The first 3D coordinate system is transformed through a three-dimensional $\boldsymbol{R}$ followed by a translation $\boldsymbol{t}$, which we will take into account as $\mathbf{O}^{(2)}\mathbf{O}^{(1)}$, to obtain the second coordinate system. The coordinate systems and transformation are right-handed.

The rotation have to satisfy

$$\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{R}^\top\boldsymbol{R} = \mathbf{I}_3 \quad \boldsymbol{R} = [\,\mathbf{r}_1\,\mathbf{r}_2\,\mathbf{r}_3\,] \text{ with } \mathbf{r}_i^\top\mathbf{r}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad i,j = 1,2,3$$

which gives six independent quadratic constaints on the three angles. However, in order to restrict the movements to a 3D space we must also have

$$\det\boldsymbol{R} = 1$$

that is, $\boldsymbol{R}$ must be a *proper* orthonormal matrix.

Taking the first coordinate system as $\mathbf{A} = \mathbf{I}_3$ and the rotation $\mathbf{B} = \boldsymbol{R}$, this change of coordinates in 3D is

$$\mathbf{y}^{[\mathbf{B}]} = \boldsymbol{R}\mathbf{y}^{[\mathbf{A}]}$$

where we project it in the *second* coordinate system.

The translation and the three-dimensional point $\mathbf{P}$, have the relation in the second coordinate system

$$\mathbf{O}^{(2)}\mathbf{P} = \mathbf{O}^{(2)}\mathbf{O}^{(1)} + \mathbf{O}^{(1)}\mathbf{P}$$

or in coordinates

$$\mathbf{y}_o^{(2)} = \boldsymbol{t} + \boldsymbol{R}\mathbf{y}_o^{(1)}$$

Each point the first coordinate system corresponds to the correct point in the second coordinate system, so only inliers are considered.

The measurements are

$$\mathbf{y}^{(k)} = \mathbf{y}_o^{(k)} + \delta\mathbf{y}^{(k)} \quad \delta\mathbf{y}^{(k)} \sim GI(0,\ \sigma^2\mathbf{I}_3) \quad k = 1, 2 \qquad \mathrm{E}[\delta\mathbf{y}^{(1)^\top}\delta\mathbf{y}^{(2)}] = 0$$

and the noise in the two coordinate systems are independent.

In the general case the estimation also have a scale change c

$$[\,\hat{c},\ \hat{\boldsymbol{t}},\ \hat{\boldsymbol{R}}\,] = \arg\min_{c,\boldsymbol{t},\boldsymbol{R}} \sum_{i=1}^{n} ||\mathbf{y}_i^{(2)} - c\boldsymbol{R}\mathbf{y}_i^{(1)} - \boldsymbol{t}||_2^2$$

subject to

$$\mathbf{y}_{io}^{(2)} = c\boldsymbol{R}\mathbf{y}_{io}^{(1)} + \boldsymbol{t} \qquad i = 1, \ldots,\ n$$

and the rotational constraints from above. This is a total least squares (TLS) estimation problem with the rotation depending nonlinearly on the six-dimensional $\{\mathbf{y}_i^{(1)},\ \mathbf{y}_i^{(2)}\}$ input.

We start by eliminating the translation from the objective function. In each of the two frames the average is

$$\bar{\mathbf{y}}^{(k)} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i^{(k)} \qquad k = 1, 2$$

and constraint on the averages between the two frames is

$$\bar{\mathbf{y}}^{(2)} = \hat{c}\hat{\boldsymbol{R}}\bar{\mathbf{y}}^{(1)} + \hat{\boldsymbol{t}}$$

where we assumed separability of the three estimates. So, if we already know $\hat{c}$, $\hat{\boldsymbol{R}}$, the translation can be found

$$\hat{\boldsymbol{t}} = \bar{\mathbf{y}}^{(2)} - \hat{c}\hat{\boldsymbol{R}}\bar{\mathbf{y}}^{(1)}$$

Next we estimate the scale between the two frames. The centered input variables are

$$\tilde{\mathbf{y}}_i^{(k)} = \mathbf{y}_i^{(k)} - \bar{\mathbf{y}}^{(k)} \qquad k = 1, 2$$

and the objective function becomes after subtituting $\hat{\boldsymbol{t}}$

$$J = \sum_{i=1}^{n} ||\tilde{\mathbf{y}}_i^{(2)} - c\boldsymbol{R}\tilde{\mathbf{y}}_i^{(1)}||^2 = \sum_{i=1}^{n} ||\tilde{\mathbf{y}}_i^{(2)}||^2 - 2c\sum_{i=1}^{n} \tilde{\mathbf{y}}_i^{(1\top)}\mathbf{R}^\top\tilde{\mathbf{y}}_i^{(2)} + c^2\sum_{i=1}^{n} ||\boldsymbol{R}\tilde{\mathbf{y}}_i^{(1)}||^2$$

because the middle term is a scalar and can be written in two ways. The zero mean sample variances are

$$\hat{\sigma}^{2(k)} = \frac{\sum_{i=1}^{n} ||\tilde{\mathbf{y}}_i^{(k)}||^2}{n-1} \qquad k = 1, 2$$

The rotation matrix in an orthonormal matrix with norm one

$$||\boldsymbol{R}\tilde{\mathbf{y}}_i^{(1)}|| = ||\tilde{\mathbf{y}}_i^{(1)}||$$

The sample covariance between the two frames is a symmetric positive semidefinite matrix

$$\mathbf{K} = \frac{1}{n-1}\sum_{i=1}^{n} \tilde{\mathbf{y}}_i^{(2)}\tilde{\mathbf{y}}_i^{(1)\top} = \frac{1}{n-1}\sum_{i=1}^{n} \tilde{\mathbf{y}}_i^{(1)}\tilde{\mathbf{y}}_i^{(2)\top}$$

and the objective function becomes

$$J = (n-1)\hat{\sigma}_2^2 - 2c(n-1)trace[\boldsymbol{R}^\top\mathbf{K}] + c^2(n-1)\hat{\sigma}_1^2$$

This is a quadratic equation in c and the solution is

$$\frac{dJ}{dc} = 0 \quad \text{which is a minimum since} \quad \frac{d^2J}{dc^2} = 2(n-1)\hat{\sigma}_1^2 > 0$$

Thus, if we find $\hat{\boldsymbol{R}}$, the scale will be also known

$$\hat{c} = \frac{trace[\hat{\boldsymbol{R}}^\top\mathbf{K}]}{\hat{\sigma}_1^2}$$

The objective function is further simplified

$$\arg\min_{\boldsymbol{R}} J = \arg\max_{\boldsymbol{R}}(trace[\boldsymbol{R}^\top\mathbf{K}])$$

A theorem... to help us.
The Cauchy-Schwarz inequality can be written for two $n \times n$ matrices

$$trace[\mathbf{A}^\top\mathbf{B}] \leq ||\mathbf{A}||\ ||\mathbf{B}||$$

where the norm can be 2-norms or Frobenius norms. We will use the Frobenius norm. Equality appears when $\mathbf{A} = \beta\mathbf{B}$ and without loss of generality (w.l.g.) we can take $\beta = 1$.
Assume $\mathbf{A} = \mathbf{B}\mathbf{Q}^\top$ where $\mathbf{Q}$ is an orthonormal matrix $\mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_n$. Then, due to the orthonormality of $\mathbf{Q}$, the inequality is

$$trace[\mathbf{Q}\mathbf{B}^\top\mathbf{B}] \leq ||\mathbf{B}||_F^2 = trace[\mathbf{B}^\top\mathbf{B}]$$

If $\mathbf{B}\mathbf{Q}^\top = \mathbf{B} = \mathbf{A}$, that is $\mathbf{Q} = \mathbf{I}_n$, the inequality reaches the upper bound and becomes an equality. The $\mathbf{D} = \mathbf{B}^\top\mathbf{B}$ is always a symmetric positive semidefinite matrix. We did prove

$$trace\mathbf{D} \geq trace[\mathbf{Q}\mathbf{D}]$$

Coming back to the objective function, the svd of the sample covariance is $\mathbf{K} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ and we can write (after substituting the rotation with the estimated rotation)

$$trace[\hat{\boldsymbol{R}}^\top\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top] = trace[\mathbf{V}^\top\hat{\boldsymbol{R}}^\top\mathbf{U}\mathbf{\Sigma}] = trace[\mathbf{Q}\mathbf{D}]$$

where $\mathbf{Q} = \mathbf{V}^\top\hat{\boldsymbol{R}}^\top\mathbf{U}$ is an orthonormal matrix and $\mathbf{D} = \mathbf{\Sigma}$ is a symmetric positive semidefinite (diagonal) matrix. A upper bound (the maximum) is reached when

$$\mathbf{Q} = \mathbf{V}^\top\hat{\boldsymbol{R}}^\top\mathbf{U} = \mathbf{I}_3 \quad \text{or} \quad \hat{\boldsymbol{R}} = \mathbf{U}\mathbf{V}^\top$$

But still we did not take into account that the rotation has to be a proper orthonormal matrix. The proper matrix have to be

$$\hat{\boldsymbol{R}} = \mathbf{U}\begin{pmatrix} 1 & & & \\ & \cdots & & \\ & & 1 & \\ & & & det(\mathbf{U}\mathbf{V}^\top) \end{pmatrix}\mathbf{V}^\top = \mathbf{U}\mathbf{Z}\mathbf{V}^\top$$

where the determinant $\det(\mathbf{U}\mathbf{V}^\top)$, which has the value 1 or -1, is taken accordingly so $\det\hat{\boldsymbol{R}} = 1$.

The scale estimate, $\hat{c}$ can be rewritten

$$\hat{c} = \frac{trace[(\mathbf{U}\mathbf{Z}\mathbf{V}^\top)^\top\ \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top]}{\hat{\sigma}_1^2} = \frac{trace[\mathbf{\Sigma}\mathbf{Z}]}{\hat{\sigma}_1^2}$$

in the second coordinated system. The physical meaning of the (inlier) similarity transformation can be seen if we rearrange the changing part of the objective function

$$c\,trace[\boldsymbol{R}^\top \mathbf{K}] = c\,trace[\mathbf{K}\boldsymbol{R}] = c\,trace[\boldsymbol{R}\mathbf{K}]$$

where $\mathbf{K}$ is symmetric and positive semidefinite. Than

$$\frac{1}{n-1}\sum_{i=1}^{n} c\,\mathbf{R}\tilde{\mathbf{y}}_i^{(1)}\tilde{\mathbf{y}}_i^{(2\top)} = \sum_{i=1}^{n}(c\,\mathbf{R}\tilde{\mathbf{y}}_i^{(1)})\ (\tilde{\mathbf{y}}_i^{(2\top)})$$

Points in the first coordinate system are transformed into the second coordinate system. The scale is taken into account and maximum alignment is sought.

In $p > 3$ dimensions, given two $n \times p$ matrices $\mathbf{A},\ \mathbf{B}$ we want to find an orthonormal $p \times p$ matrix $\mathbf{Q}$ such that

$$\min_{\mathbf{Q}} ||\mathbf{A} - \mathbf{B}\mathbf{Q}||_F \quad \text{subject} \quad \mathbf{Q}\mathbf{Q}^\top = \mathbf{Q}^\top\mathbf{Q} = \mathbf{I}_p \quad \det\mathbf{Q} = 1$$

The proper rotation that transforms $\mathbf{B}$ to $\mathbf{A}$ is known in the literature as "orthogonal Procrutes problem". The solution (like above) is

$$\max_{\mathbf{Q}}(trace[\mathbf{Q}^\top\mathbf{B}^\top\mathbf{A}]) = \max_{\mathbf{Q}}(trace[\mathbf{Q}^\top(\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top)]) = \max_{\mathbf{Q}}(trace[(\mathbf{V}^\top\mathbf{Q}^\top\mathbf{U})\boldsymbol{\Sigma}])$$

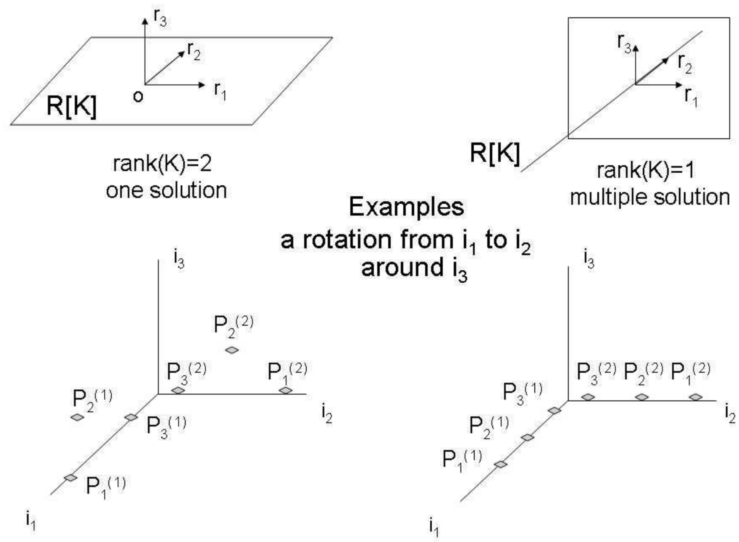
from where the $p \times p$ matrix in $\hat{\mathbf{Q}} = \mathbf{U}\mathbf{Z}\mathbf{V}^\top$ with $\mathbf{Z}$ having ones on the diagonal and $\det(\mathbf{U}\mathbf{V}^\top)$ on the last row.

Conditions for $\hat{\boldsymbol{R}}$ to be unique. Dimension of the space equal three.

If the covariance matrix $\mathbf{K}$ is full rank three, the rotation also will be full rank. In general this is the case.

If the rank of the matrix $\mathbf{K}$ is *two*, the third coordinate of the rotation, e.g., $\mathbf{r}_3$, *still* can be recovered due to the orthonormality.
Example. The rotation is around $\mathbf{i}_3$ and the three points, $\mathbf{P}_1,\ \mathbf{P}_2,\ \mathbf{P}_3$, move from the plane $(\mathbf{i}_1,\ \mathbf{i}_3)$ to the plane $(\mathbf{i}_2,\ \mathbf{i}_3)$. In this case, no line beside the rotation axis $\mathbf{i}_3$ is satisfactory.

If the rank of the matrix $\mathbf{K}$ is *one*, only a coordinate of the rotation, e.g. $\mathbf{r}_2$, is defined. In this case, there are multiple solutions for the rotation.
Example. The three points are now on the $\mathbf{i}_1$ axis and after the rotation move to the $\mathbf{i}_2$ axis. Any line through the origin and the bisector plane is satisfactory.



In general, only $\operatorname{rank}\mathbf{K} = p,\ p-1$ give unique solutions to orthogonal Procrutes problem.

Pose of an object in the coordinate system of the camera.

Three 3D vectors, $\mathbf{m}_1,\ \mathbf{m}_2,\ \mathbf{m}_3$, are given, which are independent but not orthonormal. We have to find the rotation relative to $\mathbf{I}_3$ such that

$$\min_{\mathbf{R}} \sum_{i=1}^{3} ||\mathbf{m}_i - \mathbf{r}_i|| = \min_{\mathbf{R}} ||\mathbf{M} - \mathbf{R}||_F^2$$

Developing the sum

$$||\mathbf{M} - \mathbf{R}||_F^2 = ||\mathbf{M}||_F^2 - 2\, trace[\mathbf{R}^\top \mathbf{M}] + ||\mathbf{R}||_F^2$$

but $||\mathbf{M}||_F^2$ is a constant of the problem and $||\mathbf{R}||_F^2$ is equal to three. The rotation matrix is found

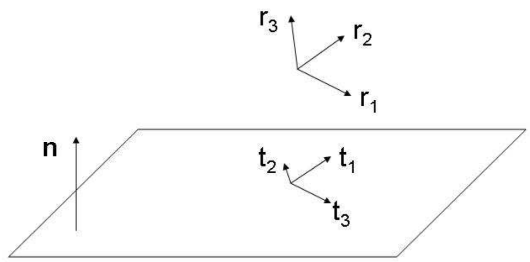
$$\max_{\mathbf{R}} trace[\mathbf{R}^\top \mathbf{M}]$$

as before... Note we have a unique solution even when one of the $\mathbf{m}$ vectors is unknown.

Orthogonal frame reconstruction.

Appears in 3D motion problems and object recognition. Given three independent vectors in *a plane* $\mathbf{T} = [\ \mathbf{t}_1\ \mathbf{t}_2\ \mathbf{t}_3\ ]$, which represent the 2D projection of a 3D

Othogonal frame reconstruction



Given three noncollinear vectors $t_{1..3}$ in a plane, find the normal vector **n** and the three dimensional frame **R**

rotation matrix $\boldsymbol{R}$

$$\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{R}^\top\boldsymbol{R} = \mathbf{I}_3 \qquad \det\boldsymbol{R} = \pm 1$$

find the rotation matrix and the unit normal to the plane $\mathbf{n}$.

The unit normal $\mathbf{n}$ is perpendicular to the plane which is the range of $\mathbf{T}$. The svd of $\mathbf{T}$ is rank two. The projection in the range of $\mathbf{T}$ is

$$\mathbf{P} = \mathbf{I}_3 - \mathbf{n}\mathbf{n}^\top = \mathbf{u}_1\mathbf{u}_1^\top + \mathbf{u}_2\mathbf{u}_2^\top$$

from where $\hat{\mathbf{n}}$ is easy to recover, $\hat{\mathbf{n}} = \mathbf{u}_3$.

The 3D rotation matrix is the minimum in 2D of

$$\arg\min_{\boldsymbol{R}} ||\mathbf{T} - \mathbf{P}\boldsymbol{R}||_F^2 = \arg\max_{\boldsymbol{R}} trace[\boldsymbol{R}^\top\mathbf{P}^\top\mathbf{T}] = \arg\max_{\boldsymbol{R}} trace[\boldsymbol{R}^\top\mathbf{T}] = \cdots$$

because $\mathbf{P}^\top\mathbf{T} = \mathbf{T}$ and the solution is like before.

The matrix $\mathbf{Q} = \mathbf{I}_3 - 2\mathbf{n}\mathbf{n}^\top$ is the reflection of the projection

$$\mathbf{P}\mathbf{Q} = \mathbf{I}_3 - 3\mathbf{n}\mathbf{n}^\top + 2\mathbf{n}\mathbf{n}^\top\mathbf{n}\mathbf{n}^\top = \mathbf{P}$$

If for *one* 3D vector we know a priori the orientation (away/toward the image plane), we have a unique solution.