Orthognal Procrustes Problem

Consider where have a set of points $\mathbf{P}_1 \in \mathbb{R}^{2\times N}$ and another set of points $\mathbf{P}_2 \in \mathbb{R}^{2\times N}$ such that $\mathbf{P}_1$ and $\mathbf{P}_2$ are related by an orthonormal transformation $\mathbf{R}$ such that $\mathbf{P}_1 = \mathbf{R}\mathbf{P}_2 + \mathbf{E}$ where $\mathbf{E} \in \mathbb{R}^{2\times N}$ is an error (or noise) matrix. The aim is to find $\mathbf{R}$ given $\mathbf{P}_1$ and $\mathbf{P}_2$.
The standard least squares solution given by $\mathbf{R} = \mathbf{P_1}\mathbf{P_2}^T(\mathbf{P_2}\mathbf{P_2}^T)^{-1}$ will fail, because it will not impose the constraint that $\mathbf{R}$ is orthonormal (excepting the situation where $\mathbf{E} = \mathbf{0}$). You can try this out in MATLAB! To solve for $\mathbf{R}$, we seek to minimize the following quantity:

$$
\begin{aligned}
E(\mathbf{R}) &= \|\mathbf{P_1} - \mathbf{R}\mathbf{P_2}\|_F^2 && (1)\\
&= \text{trace}((\mathbf{P_1} - \mathbf{R}\mathbf{P_2})^T(\mathbf{P_1} - \mathbf{R}\mathbf{P_2})) && (2)\\
&= \text{trace}(\mathbf{P_1^T P_1} + \mathbf{P_2^T R^T R P_2} - \mathbf{P_2^T R^T P_1} - \mathbf{P_1^T R P_2}) && (3)\\
&= \text{trace}(\mathbf{P_1^T P_1} + \mathbf{P_2^T P_2} - \mathbf{P_2^T R^T P_1} - \mathbf{P_1^T R P_2}) && (4)\\
&= \text{trace}(\mathbf{P_1^T P_1} + \mathbf{P_2^T P_2}) - 2\text{trace}(\mathbf{P_1^T R P_2}) && (5)
\end{aligned}
$$

The second-last equality follows as $\mathbf{R^T R} = \mathbf{I}$. The last equality follows using the fact that trace of a matrix equals the trace of its transpose.

The first two terms are constant and the last term has a negative sign. Hence, minimizing $E(\mathbf{R})$ w.r.t. $\mathbf{R}$ is equivalent to maximizing $\text{trace}(\mathbf{P_1^T R P_2})$ w.r.t. $\mathbf{R}$. Now, we have

$$
\begin{aligned}
\text{trace}(\mathbf{P_1^T R P_2}) &= \text{trace}(\mathbf{R P_2 P_1^T})(\because \text{trace}(AB) = \text{trace}(BA)) && (6)\\
&= \text{trace}(\mathbf{R U' S' V'^T}) \text{ using SVD of } \mathbf{P_2 P_1^T} = \mathbf{U' S' V'^T} && (7)\\
&= \text{trace}(\mathbf{S' V'^T R U'}) = \text{trace}(\mathbf{S' X}) \text{ where } \mathbf{X} \text{ is orthonormal} && (8)\\
&= \sum_i S'_{ii} X_{ii} && (9)
\end{aligned}
$$

This expression is maximized if $\mathbf{X}_{ii} = 1$ all along its diagonal. As $\mathbf{X}$ is orthonormal, we must have $\mathbf{X} = \mathbf{I}$, and hence $\mathbf{V'^T R U'} = \mathbf{I}$ giving $\mathbf{R} = \mathbf{V'}\mathbf{U'}^T$ where $\mathbf{U'}$ and $\mathbf{V'}$ are obtained from the SVD of $\mathbf{P_2 P_1^T}$.
This is a deceptively tricky problem, as it is so very easy to fall into the 'pseudo-inverse' trap :-). However, it is an important and fundamental problem in computer vision, graphics and medical imaging (especially in a sub-branch called as 'statistical shape analysis'). It is popularly called as the 'orthogonal procrustes problem' and the first complete solution to it was first proposed by a person named Peter Schönemann in a psychology (!) journal in 1966. A popular application scene where this problem is useful is as follows: Consider N pairs of corresponding points marked out on two 3D models of a person (acquired from say a Kinect from two different positions). What is the rotation and translation in between these points? This has obvious applications in person recognition and aligning scans of humans or different types of objects.