SOUTH BANK UNIVERSITY
·LONDON·

SCHOOL OF COMPUTING, INFORMATION SYSTEMS & MATHEMATICS

Some Remarks on the Statistics of Pose Estimation

by

J.M. Selig

Technical Report SBU-CISM-00-25

Some Remarks on the Statistics of Pose Estimation

J.M. Selig
School of Computing, Information Systems & Mathematics
South Bank University
London SE1 0AA, U.K.

Abstract

This work argues that directional statistics are important in Robotics. That is, statistics on general manifolds. Historically the subject began with statistics on circles and spheres, hence the title. It is still a relatively new discipline, see [4] for an overview. In order to make the case a specific example is studied: Finding the rigid transformation undergone by a camera from a knowledge of the images of a number of points. This very simple, perhaps naïve, example allows us to study differnet models for error in the observations and estimators for the rigid transformation.

1 Introduction

A large number of important problems in Robotics, both in Robot vision and kinematics, seem to come down to finding a rigid body transformation given some data describing its effects on points or lines and so forth.

The most obvious examples are in robot vision where a common problem is to compute the motion of the camera given a number of point correspondences in successive images. For mobile robots we need to find the position and orientation of the robot given data from a number of different sensors, perhaps a camera or a sonar range-finder or several other possible sensors. An interesting problem for industrial robots is the so-called sensor calibration problem (this is not related to camera calibration familiar in robot vision). Here the robot has a camera attached to its end-effector, by moving the end-effector through known motions we capture successive images. From this information we must find the rigid change of coordinates which relates the robot's tool coordinates to the coordinate system of the camera.

In many cases the geometrical problem is quite simple if the data is accurate. In reality the data comes from measurements and all measurements are subject to noise. The real question that these geometrical problems pose is thus: What is the best estimate of the rigid transformation given noisy data?

Many workers have proposed solutions to various problems over the years, however, these solutions do not take into account the geometry of the group of rigid transformations. For example, it is fairly common find least squares solutions for the matrix elements subject to the constraints that the matrix is a group element. The difficulty here is that there are many ways to embed the group of proper rigid motions into Euclidean spaces of various dimensions. The results are different for different embeddings, this is clear because the group does not have a bi-invariant Euclidean metric so none of the embeddings can be isometric.

A more natural approach would be to work on the group of rigid body motions itself. This would ensure that results would have physical (coordinate-free) meaning. There are however, many difficult technical problems to overcome.

In this work we study some of these difficulties by looking at a representative problem. In this problem a calibrated camera observers a set of points whose positions are known. The idea is to infer the position and orientation (the configuration or pose) of the camera from the observed positions of the points.

After describing this problem in a little more detail several different models for the errors in the data are studied. Next some different estimators for the pose are derived, for example the different error models will give different maximum likelihood estimators.

Here we will assume that we know the position in space of a number of points $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$. Now the camera is subjected to a rigid transformation X^{-1} or equivalently the points are subject to a transformation X. After the transformation the points are observed by the camera, but the camera can only observe a line joining its centre of projection to a point. We denote the corresponding lines observed by the camera by $L_1, L_2, \ldots, L_n$. If these lines were observed with no error then the transformed point $X(\mathbf{b}_i)$ would lie on the line L_i. Of course there will be errors, so this will not be true in general.

2 Error Models

In this section we will look at several different possible models for errors. These different models are distinguished mainly by the assumptions made as to were the errors occur. That is, whether the errors occur in the image plane of the vision system or in space in the coordinates of the observed points. In order to make reasonable comparisons, however, we must consider the error distributions on the same space. A reasonable, space to use is the projective space of lines through the centre of projection of the camera, $\mathbb{PR}^2$. Hence, in the following we will map the probability distributions to this space.

We can use spherical polar coordinates for the space of lines,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}$$

We restrict the range of θ to $0 \leq \theta \leq \pi/2$, so that each un-directed line only has one set of coordinates (except for the z-axis, $\theta = 0$, $\phi =$ any). See figure 1.

We will also need the Jacobian of this change of coordinates,

$$dx\,dy\,dz = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix} dr\,d\theta\,d\phi = r^2\sin\theta\,dr\,d\theta\,d\phi$$

2.1 Image-plane Gaussian

This seems to be the most commonly used model in computer vision, see [2] for example. Essentially, this model assume that all the errors occur in the vision system, the motion was precise but errors were made detecting the points in the image. The distribution is represented by a probability density function of the form:

$$g_I(\mathbf{x}_I; \mathbf{x}_0, K) = C_1 \exp\left\{(\mathbf{x}_I - \mathbf{x}_0)^T K (\mathbf{x}_I - \mathbf{x}_0)\right\}$$

Here the image-plane coordinates are,

$$\mathbf{x}_I = \begin{pmatrix} x_I \\ y_I \end{pmatrix} = \begin{pmatrix} xf/z \\ yf/z \end{pmatrix} = \begin{pmatrix} f\tan\theta\cos\phi \\ f\tan\theta\sin\phi \end{pmatrix}$$

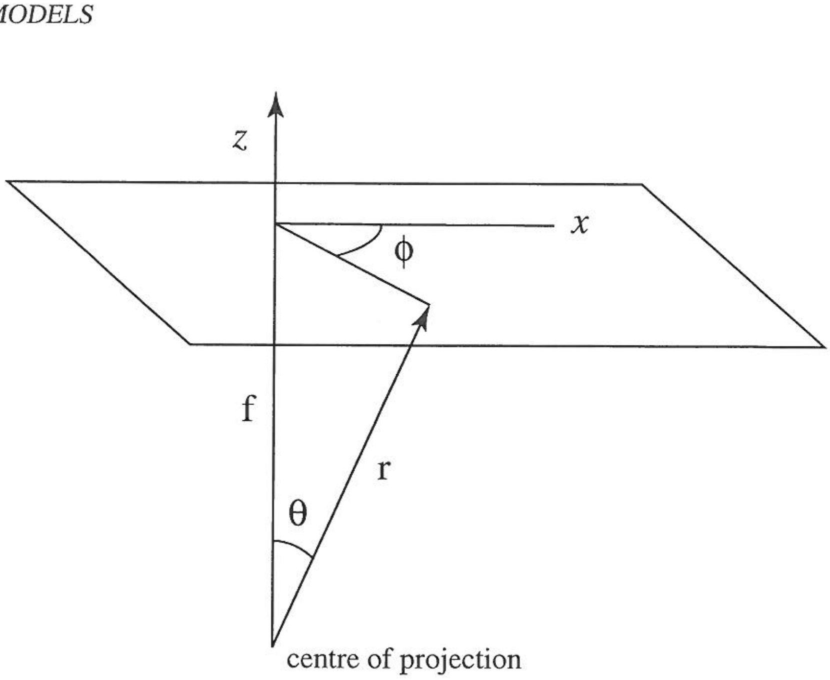

Figure 1: Spherical Polar Coordinates

$\mathbf{x}_0$ is the mode of the distribution and K is a symmetric matrix, usually called the covariance matrix. The length f specifies the distance from the image-plane to the centre of projection at the origin. Notice that C_1 has been written for the normalisation constant here. We will not be concerned with these constants in this article and will simply write them as C_i in the following.

In the spherical polar coordinates this distribution becomes,

$$g_I'(\theta,\phi;\alpha,\beta,K) = C_1 f^2 \exp\left\{(\mathbf{x}_I-\mathbf{x}_0)^T K(\mathbf{x}_I-\mathbf{x}_0)\right\}\sec^2\theta\tan\theta$$

where,

$$\mathbf{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} f\tan\alpha\cos\beta \\ f\tan\alpha\sin\beta \end{pmatrix}$$

For the simple case where $(\mathbf{x}_I-\mathbf{x}_0)^T K(\mathbf{x}_I-\mathbf{x}_0) = -\left((x_I-x_0)^2+(y_I-y_0)^2\right)/2\sigma^2$, the result becomes,

$$g_I'(\theta,\phi;\alpha,\beta,\sigma) = C_1 f^2 \exp\left\{\frac{f^2}{2\sigma^2}(\sec^2\theta+\sec^2\alpha-2\sec\theta\sec\alpha\cos\psi)\right\}\sec^2\theta\tan\theta$$

Here, $\cos\psi = \cos\theta\cos\alpha+\sin\theta\sin\alpha\cos(\phi-\beta)$, that is ψ is the angle between a general line and the line through the mode of the distribution, see figure 2.

In the language of directional statistics this is an example of a wrapped distribution. A distribution on the tangent plane of a manifold is mapped to a distribution on the manifold itself, see [4, §13.2]. Actually, rather than using the exponential map to wrap the distribution onto $\mathbb{PR}^2$ we are using the camera projection.

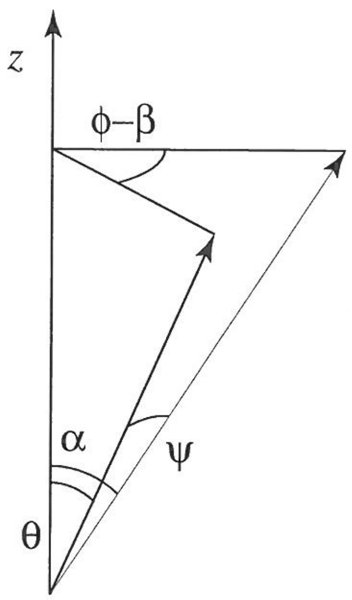

Figure 2: The Angle between Two Lines

2.2 Object-Space Gaussian

In this section suppose that the camera is fixed and the points it is observing are fixed to body that is not completely rigid. In this case it would be reasonable to assume that the positions of the points in space obey a 3-D Gaussian with a probability density function of the form,

$$g_o(x,y,z;x_0,y_0,z_0,\sigma) = C_2 \exp\{-((x-x_0)^2+(y-y_0)^2+(z-z_0)^2)/2\sigma^2\}$$

To find the density function for lines implied by the above, we integrate this function along the lines. This procedure is reminiscent of Radon transform. In the language of directional statistics this procedure is known as the projection of a distribution, [4, pp. 178–179].

$$g_o'(\theta,\phi;x_0,y_0,z_0) = \int_{-\infty}^{+\infty} g_o r^2 \sin\theta \, dr$$

To perform the integral we need to convert the above to polar coordinates,

$$g_o'(\theta,\phi;x_0,y_0,z_0,\sigma) = C_2 \sin\theta \int_{-\infty}^{+\infty} r^2 e^{(r^2+r_0^2-2rr_0\cos\psi)/2\sigma} dr$$

where we assume that the mode of the distribution has the polar coordinates,

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} r_0 \sin\alpha\cos\beta \\ r_0 \sin\alpha\sin\beta \\ r_0\cos\alpha \end{pmatrix}$$

and $\cos\psi = \cos\theta\cos\alpha + \sin\theta\sin\alpha\cos(\phi-\beta)$. Now we can substitute $R = (r-r_0)/(\sqrt{2}\sigma)$ and obtain the result,

$$g'_o(\theta,\phi;\alpha,\beta,\sigma) = \sqrt{2\pi}C_2\sigma^3 \exp\left\{-\frac{r_0^2}{\sigma^2}\sin^2\psi\right\}(1+2\frac{r_0^2}{\sigma^2}\cos^2\psi)\sin\theta$$

As above, the angle ψ is the angle between the particular line and the direction to the mode of the Gaussian distribution.

When the mode of the Gaussian distribution coincides with the centre of projection it is clear that we should expect a uniform distribution on the space of lines $\mathbb{PR}^2$. This is indeed the case, remembering that, in these coordinates the distribution function for the uniform distribution is proportional to $\sin\theta$. Of course this is not a very practical situation since, if a point were located at the centre of projection we would not be able to see it with our camera.

For a Gaussian with a more general covariance matrix the integral becomes impossible to perform in terms of elementary functions. However, good approximations do exist.

2.3 The Watson Distribution

This is a distribution defined on spheres, it has anti-podal symmetry and hence can be considered as a distribution function on $\mathbb{PR}^2$, the space of lines through the origin in $\mathbb{R}^3$.

The Watson distribution is defined by the density function,

$$W(\pm\mathbf{x};\boldsymbol{\mu},\kappa) = C_3\exp\left\{\kappa(\boldsymbol{\mu}^T\mathbf{x})^2\right\}$$

here $\pm\boldsymbol{\mu}$ is the modal direction and κ is the concentration parameter. Both $\boldsymbol{\mu}$ and $\mathbf{x}$ are intended to be unit vectors here. In spherical coordinates we can write,

$$\begin{aligned}\mathbf{x} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^T \\ \boldsymbol{\mu} &= (\sin\alpha\cos\beta, \sin\alpha\sin\beta, \cos\alpha)^T\end{aligned}$$

In these coordinates the density function becomes,

$$W(\theta,\phi;\alpha,\beta,\kappa) = C_3\exp\left(\kappa\cos^2\psi\right)\sin\theta$$

where, as in the section above, $\cos\psi = \cos\theta\cos\alpha + \sin\theta\sin\alpha\cos(\phi-\beta)$, and ψ gives the angle between the line under consideration and the modal direction.

Notice that this is very similar to the projected normal distribution we met above. If we set the distribution parameter κ, to $\kappa = (r_0^2/\sigma^2)$ then for small κ the two distributions functions become almost identical (remember that $\sin^2\psi = 1-\cos^2\psi$). However, as we saw above this is the case where the distribution function approximates a uniform distribution, so this observation is not particularly useful in practice where we expect large concentration κ or small variance σ^2.

Clearly there are many different possible distributions for lines through the origin. The question is which is the best one to choose? Of course the answer will depend on the problem at hand. However, when no other information is available the Watson distribution is probably the best one to choose. There are several reasons for this, it is by far the simplest to work with and it shares many properties with the Gaussian distribution in Euclidean space.

3 Estimators

It is well known that for regression problems in Euclidean space the least-squares estimator is the maximum likelihood estimator exactly when the data points are assumed to obey Gaussian distributions with common covariance. Here the situation if far less clear.

Some results are known, for example suppose we try to estimate the rotation which takes a set of points on the sphere to a set of measured points on the sphere. If we assume that the measured points obey a Fisher distribution then the maximum likelihood estimator is the same as the least squares estimator, see [7].

In general, all the methods considered below work in the following manner; a function is defined on the group space, either a residual or a log-likelihood function. Then the group element that minimises this function must be found. An advantage here is that since the manifold we are working on is a group space we can use the Lie algebra elements as (left-invariant) vector fields on the manifold. So if we find the derivative of the function with respect to an arbitrary element of the Lie algebra and set this to zero we get conditions for the function to be stationary. Notice that this completely avoids the need for Lagrange multipliers.

Suppose X is an element of the group of rigid body motions $SE(3)$, The effect of such a group element on a point in $\mathbb{R}^3$ is given by a rotation and a translation,

$$X(\mathbf{b}) = R\mathbf{b} + \mathbf{t}$$

Here, R is a 3×3 rotation matrix and $\mathbf{t}$ a 3-dimensional translation vector. Usually we work in a representation of the group, for example, the 4-dimensional representation where,

$$X = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix}$$

(abusing notation slightly). The advantage of this representation is that we can represent the action of the group on points by matrix multiplication,

$$X\tilde{\mathbf{b}} = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathbf{b} \\ 1 \end{pmatrix} = \begin{pmatrix} R\mathbf{b}+\mathbf{t} \\ 1 \end{pmatrix}$$

where we have embedded the 3-vector $\mathbf{b}$ into $\mathbb{R}^4$, $\tilde{\mathbf{b}}^T = (\mathbf{b}^T, 1)$, this is sometimes called the homogeneous representation.

Let $F(X)$ is a function defined on the group. To differentiate this function with respect to a Lie algebra element S we move the along the vector field defined by S, this will give a function value, $F(e^{tS}X)$. In the usual way we then subtract the original function value, divide by the parameter t and proceed to the limit $t \to 0$.

$$\partial_S F(X) = \lim_{t\to 0}\big(F(e^{tS}X) - F(g)\big)/t = \frac{d}{dt}F(e^{tS}X) = F'(X)SX$$

For more details on this technique see [6]. In general we will use a linear representation of the group and hence the Lie algebra element S must come from the corresponding representation of the Lie algebra.

3.1 Least Squares

Returning to our model problem. Suppose we transform the known points and then compute the distances between the transformed points and the corresponding measured lines. Now we seek the transformation which minimises the sum of the squares of these distances.

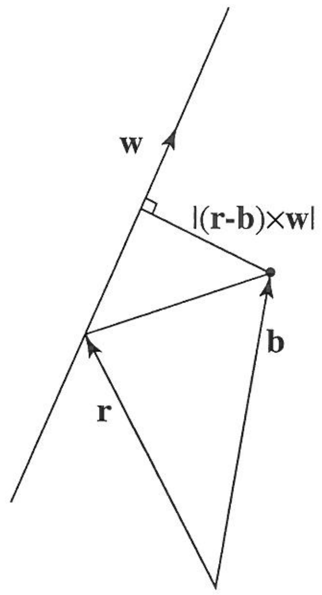

Figure 3: Shortest Distance Between a Point and a Line

The square of the distance between a point $\mathbf{b}$ and a line is given by,

$$d^2 = (\mathbf{u} - \mathbf{b} \times \mathbf{w}) \cdot (\mathbf{u} - \mathbf{b} \times \mathbf{w})$$

Here, $\mathbf{w}$ is a unit vector in the direction of the line and $\mathbf{u}$ is the moment of the line, $\mathbf{u} = \mathbf{r} \times \mathbf{w}$ with $\mathbf{r}$ and point on the line. See figure 3. The sum of the squares of the distances between the points and their lines can be written as,

$$D(X) = \sum_i (L_i X \tilde{\mathbf{b}}_i)^T (L_i X \tilde{\mathbf{b}}_i)$$

The lines are represented here by 3×4 matrices with the partitioned form,

$$L = \begin{pmatrix} W & \mathbf{u} \end{pmatrix}$$

with W the anti-symmetric matrix such that $W\mathbf{x} = \mathbf{w} \times \mathbf{x}$ for any vector $\mathbf{x}$. Now we can minimise this over the group by differentiating along an arbitrary left invariant vector field S,

$$\partial_S D(X) = 2 \sum_i (L_i X \tilde{\mathbf{b}}_i)^T (L_i S X \tilde{\mathbf{b}}_i)$$

see above. Setting this equal to zero for arbitrary S produces the equations,

$$\sum_i \left(W_i^T W_i (R\mathbf{b}_i + \mathbf{t}) + W_i^T \mathbf{u}_i\right) \times (R\mathbf{b}_i + \mathbf{t}) = 0$$

and

$$\sum_i \left(W_i^T W_i (R\mathbf{b}_i + \mathbf{t}) + W_i^T \mathbf{u}_i\right) = 0$$

The estimate is the rotation R and translation $\mathbf{t}$ which solves these equations. Certainly, if there are no errors in the data this estimator will give the rigid transformation precisely.

It is difficult to see how these equations could be solved symbolically. There are a few simplification which may help, for example, since all the lines will pass through the camera's centre of projection, we can take this point as the origin and then all the $\mathbf{u}_i$s will vanish. We can make the equations a bit neater by writing $\mathbf{p}_i = R\mathbf{b}_i + \mathbf{t}$ and using the fact that $W_i\mathbf{p}_i = \mathbf{w}_i \times \mathbf{p}_i$ and various vector identities to get:

$$\sum_i (\mathbf{p}_i \cdot \mathbf{w}_i)(\mathbf{p}_i \times \mathbf{w}_i) = 0 \tag{1}$$

and

$$\sum_i \mathbf{w}_i \times (\mathbf{p}_i \times \mathbf{w}_i) = 0 \tag{2}$$

However, this does not seem to help us untangle the rotational from the translational part of the problem.

Since we have not assumed a distribution for the errors here there are several questions we could legitimately ask. For example, for which error distribution does this give the maximum likelihood estimator? Also, what are the statistics of this estimator? That is, given a distribution for the lines, how is the estimate distributed on the space of rigid body transformations? This last question can be asked for the maximum likelihood estimators derived below.

3.2 Maximum Likelihood for Lines

Here assume that the errors are distributed according to a Watson distribution with known concentration parameter κ. The data we have consists of pairs, points $\mathbf{b}_i$ and the corresponding directions $\mathbf{w}_i$. The log-likelihood of the data is thus,

$$K(X) = \kappa \sum_i (R\mathbf{b}_i + \mathbf{t})^T \mathbf{w}_i \mathbf{w}_i^T (R\mathbf{b}_i + \mathbf{t})/|R\mathbf{b}_i + \mathbf{t}|^2 + n \ln C_3$$

Again this can be considered as a function of the group of rigid transformations. Hence, to minimise it we must differentiate with respect to arbitrary Lie algebra element and set the result to zero.

To facilitate this let us write, $\mathbf{p}_i = R\mathbf{b}_i + \mathbf{t}$ as above and $\hat{\mathbf{p}}_i$ for the unit vector in this direction, $\hat{\mathbf{p}}_i = \mathbf{p}_i/|\mathbf{p}_i|$. Now since we can write $\hat{\mathbf{p}}_i = (\mathbf{p}_i^T \mathbf{p}_i)^{-1/2} \mathbf{p}_i$, we can calculate that,

$$\partial_S \hat{\mathbf{p}}_i = \frac{1}{|\mathbf{p}_i|^3} \left((\mathbf{p}_i^T \mathbf{p}_i) I_3 - (\mathbf{p}_i \mathbf{p}_i^T)\right) \partial_S \mathbf{p}_i$$

The term $\partial_S \mathbf{p}_i$ is given by,

$$\partial_S \mathbf{p}_i = \omega \times \mathbf{p}_i + \mathbf{v} = (P_i^T, I_3) \begin{pmatrix} \omega \\ \mathbf{v} \end{pmatrix}$$

where ω and $\mathbf{v}$ are the first three and second three components of an arbitrary Lie algebra element written in the adjoint representation—a 6-dimensional representation. The 3×3 identity matrix is I_3 and P_i represents the 3×3 anti-symmetric matrix corresponding to $\mathbf{p}_i$, $P_i\mathbf{x} = \mathbf{p}_i \times \mathbf{x}$ for any vector $\mathbf{x}$.

Putting all of this together we have,

$$\partial_S K(X) = 2\kappa \sum_i \hat{\mathbf{p}}_i^T \mathbf{w}_i \mathbf{w}_i^T \frac{1}{|\mathbf{p}_i|^3} \left((\mathbf{p}_i^T \mathbf{p}_i) I_3 - (\mathbf{p}_i \mathbf{p}_i^T)\right) (P_i^T, I_3) \begin{pmatrix} \omega \\ \mathbf{v} \end{pmatrix}$$

Setting this to zero for arbitrary ω and $\mathbf{v}$ gives two vector equations,

$$\sum_i \frac{1}{|\mathbf{p}_i|^2}(\mathbf{p}_i \cdot \mathbf{w}_i)(\mathbf{p}_i \times \mathbf{w}_i) = 0 \tag{3}$$

and

$$\sum_i \frac{1}{|\mathbf{p}_i|^4}(\mathbf{p}_i \cdot \mathbf{w}_i)(\mathbf{p}_i \times (\mathbf{w}_i \times \mathbf{p}_i)) = 0 \tag{4}$$

3.3 Maximum Likelihood in the Image-plane

In this section we look at the usual case where the errors are given by Gaussians in the image plane. The log-likelihood function can be written as,

$$N(X) = \frac{-1}{2\sigma^2}\sum_i \Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^2 + n\ln C_1$$

where the projection map pr is given by,

$$\mathrm{pr}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} fx/z \\ fy/z \end{pmatrix}$$

As usual we must differentiate this and set the result to zero, in the previous two examples we could perform the differentiation with respect to an arbitrary Lie algebra element and produce a pair of 3-vector equations. In this case this does not seem to be possible. Hence, rather than use an arbitrary Lie algebra element we can take a basis for the Lie algebra, necessarily consisting of six-algebra independent elements.

Let use write,

$$\mathbf{p}_i = R\mathbf{b}_i + \mathbf{t} = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

So that,

$$\partial_S\,\mathrm{pr}\begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} f\frac{\partial_S(x)z - x\partial_S(z)}{z^2} \\ \frac{\partial_S(y)z - y\partial_S(z)}{z^2} \end{pmatrix}$$

Now as our six basis vectors for the Lie algebra we take the infinitessimal rotations about the coordinate axes and the infinitessimal translations parallel to the coordinate axes. The six equations for maximum liklihood are then,

$$\Sigma_i \tfrac{f}{z_i^2}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} x_i y_i \\ y_i^2 + z_i^2 \end{pmatrix} = 0, \quad \Sigma_i \tfrac{f}{z_i}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0,$$

$$\Sigma_i \tfrac{f}{z_i^2}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} x_i^2 + z_i^2 \\ x_i y_i \end{pmatrix} = 0, \quad \Sigma_i \tfrac{f}{z_i}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0,$$

$$\Sigma_i \tfrac{f}{z_i}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} -y_i \\ x_i \end{pmatrix} = 0, \quad \Sigma_i \tfrac{f}{z_i^2}\Big(\mathrm{pr}(\mathbf{p}_i) - \mathrm{pr}(\mathbf{w}_i)\Big)^T \begin{pmatrix} x_i \\ y_i \end{pmatrix} = 0.$$

4 Concluding Remarks

In the above we have assumed that the parameters (covariance or concentration) of the distributions are known. It is more usual to use the data to estimate these quantities also. To do this we need to know explicitly how the normalisation constant of the distribution C_i here, depend on the parameters. Then we can differentiate the log-likelihood with respect to the parameters, this is not always straightforward since the the normalisation constant is usually some hypergeometric function of the parameter.

As mentioned above,we need to solve the equations we have derived to find our best estimate. This will usually involve some numerical technique. Several approaches to this exist in the literature see for example [3]. However, one technique which suggests itself from the work above is steepest decent. In the first two cases we were able to compute the gradient of the function to be minimised, recall that,

$$dF(S) = \partial_S F$$

Now it is relatively straightforward to find the Lie algebra element S which maximises $-dF$ and then we can move a small amount in the group by multiplying our current estimate by the group element $e^{\delta S}$ for some small step length δ.

An alternative approach might be to turn this into a dynamical system. That is we consider $F(X)$ as a potential energy function then assume a kinetic energy function by choosing an inertia tensor and we would probably want some damping in the system too. Now we simply use standard numerical techniques to integrate the equations of motion.

It is "well known" that many of these estimators produce biased results. However, there is a real difficulty in understanding what bias means in this context. Classically an estimator is unbiased if the mean of its results coincides with the actual value we are trying to estimate. In the problems we are considering the object we are trying to estimate is an element of a Lie group and there does not seem to be a natural way to extend the idea of a mean value to this space. However, there is work on extending some of these ideas to more general spaces. In [5] the idea of mean points for distributions on Riemannian manifolds are defined and an intrinsic definition of unbiased estimators in such situations is given. Notice that it seem reasonable to require that any result we compute should be independent of the coordinate frame chosen.

We have not mentioned Bayesian estimators so far. As usual to use a Bayesian estimator we need a prior distribution. In this case the prior distribution will be a distribution on the group space. In modern Bayesian theory much is made of maximum entropy priors, these distributions can be thought of as representing ones state of knowledge or ignorance before the experiment is performed. Unfortunately there does not seem to be a good definition of entropy for general manifolds. In particular the group space for the group of rigid body motions is a non-compact manifold. On the circle there is a definition of entropy and with suitable assumptions the maximum entropy distributions turn out to the the von Mises distribution, see [4, p.42].

So for Bayesian estimation, several reasons we are led to the problem of trying to understand distributions on the group of rigid body transformation. As far as the author is aware there is almost no previous literature on this. However, some obvious examples suggest themselves. First of all, if we choose an origin for our coordinates we can split the group into the semi-direct product of rotations about the origin with the translations, $SE(3) = SO(3) \rtimes \mathbb{R}^3$. Now we can choose a product measure, that is a distribution on the rotation and one on the translations. An obvious choice for the translations might be a Gaussian. The rotation group $SO(3)$ is know to have a group space isomorphic to the projective space $\mathbb{PR}^3$ and hence we might choose a Watson distribution here. Another approach, which does not need a origin to be chosen beforehand is the following. Suppose we have a potential function defined on the group space, for example imagine a rigid body suspended by a number of springs. In a simplified model of this situation it can be shown that the potential energy has a single minimum. Now suppose we take the exponential of the negative of the potential. This function can be integrated of the group in principle and hence a normalising constant will exist to make this into a probability distribution on the group. The result will be reminiscent of the Bingham distributions for spheres. Note that the stable equilibrium of the rigid body and springs becomes the mode of the distribution.

In camera self-calibration the task is to find the parameters of the vision system from sets of point correspondences. The set of camera parameters can be conveniently expressed as element of the Lie group $SL(3)$. In other problems, there is no unique solution, even when the data is precise. Rather these problems are incompletely specified. An example would be trying to find a rigid motion given sets of point correspondences between successive images. The problem is that, if we don't know how far the points are from the camera we cannot distinguish between a small rotation of nearby points and a larger rotation of more distant points. A common way around this difficulty is to assume that the points are all a unit distance from the camera's centre of projection. A better way of thinking about this is to quotient out the ambiguity. That is, take the space of rigid body transformations quotient by the equivalence relation which identifies two transformations if they produce the same data. Now we have a well defined problem on the quotient space.

In summary, we should keep an open mind about which Lie group we are using and generally consider the problem of estimation on general Lie groups and their quotients.

Acknowlegements

Many thanks to the organisers of AFPAC 2000 for inviting me to present this work. In particular to G. Sommer for introducing me to the reference [1].

References

[1] A. Chiuso and G. Picci Visual Tracking of Points as Estimation ion the Unit Sphere. In *Lecture Notes in Control and Information Sciences*, **237**. Springer Verlag, 1998.

[2] K. Kanatani. *Geometric Computation for Machine Vision*. Clarendon Press, Oxford, 1993.

[3] Lasenby, J. and Lasenby, A.N. *Estimating tensors for matching over multiple views.* Phil.Trans.Roy.Soc.-A , 356, 1267-1282. Ed. J. Lasenby, A. Zisserman, H.C. Longuet-Higgins and R. Cipolla. 1998.

[4] K.V. Mardia and P.E. Jupp. *Directional Statistics* 2nd edition. John Wiley, Chichester, 1999.

[5] J.M. Oller and J.M. Corcuera. Intrinsic analysis of statistical estimation. *The Annals of Statistics*, **23**:5 pp.1562–1581,1995.

[6] J.M. Selig. Three Problems in Robotics. *Proceedings of A symposium Commemorating the Legacy, Works, and Life of Sir Robert Stawell Ball Upon the 100th Anniversary of A Treatise on the Theory of Screws*, Cambridge paper no. 8, 2000.

[7] G.S. Watson. Statistics of Rotations, In *Lecture Notes in Mathematics*, Springer Verlag, **1379**, pp.398-413, 1989.