Rigid Body Inertia Estimation in Torque Free Motion

Nick Colonnese

Spring 2011

1 Background

This paper concerns estimating the inertia matrix of an object undergoing torque free motion (imagine a satellite or asteroid). An inertia matrix J describes the mass distribution of a rigid body and has several properties: it is symmetric, positive definite, and has the property that the sum of its two smallest eigenvalues must be equal to or larger than its largest eigenvalue [Gre88].

It is possible to track features on a rigid body from which attitude and attitude rates can be obtained. Possible ways in which this is accomplished is using a pair of stereo cameras [FL06], or using range data [Zam96]. With this information there exists a method to estimate the inertia matrix of the body to a scale factor [SR09]. A scale factor is the best you can do because bodies that are high momentum and high inertia are indistinguishable from low momentum low inertia ones. The method uses the fact that the angular momentum of a body in torque free motion as observed in an inertial frame does not change. This method does not take advantage of all the known properties of J, namely that J is positive definite and that the sum of its two smaller eigenvalues must be larger than or equal to its largest eigenvalue.

This paper will describe how to include the additional information of J into estimation procedure.

2 Approach

Perhaps surprisingly the properties of J describe a convex set. Unfortunately, the constraints on J cannot be included in a straightforward way to the inertia estimation algorithm from previous research because it is a non-convex problem. The problem formulation taking advantage of the properties of of J takes the takes the form

$$\begin{array}{ll} \text{minimize} & \|Ax\|_2^2 \\ \text{subject to} & \|x\|_2 = 1 \\ & J \succ 0 \\ & \lambda_1(J) \leq \lambda_2(J) + \lambda_3(J) \end{array} \tag{1}$$

where $x \in \mathbf{R}^9$ is the variable. The first 6 entries of x correspond to the unique entries of $J \in \mathbf{R}^{3\times3}$ while the last 3 correspond to the angular momentum. $A \in \mathbf{R}^{3M\times9}$ where M is the number of measurements. $\lambda_i(J)$ represents the i^{th} largest eigenvalue of J.

$$J = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix}$$

The constraints on J are convex because the sum of the largest (smallest) eigenvalues of a symmetric matrix is convex (concave) [BV04]. The equality constraint is non-convex. This problem formulation is easily solved without the constraints on J, however with them it cannot be solved directly.

To solve this problem three different approaches are considered. The first approach is to ignore the constraints on J and solve with the singular value decomposition (SVD). The second approach is to project the SVD estimate for the inertia matrix onto space of valid inertia matrices. The third approach is to include the constraints on J directly in the estimation by linearizing about a current point to make the problem convex and then iterating. Each approach is detailed in this section.

2.1 SVD

If the constraints on J are removed from (1) then the inertia estimation problem takes the form

$$\begin{array}{ll} \text{minimize} & \|Ax\|_2^2 \\ \text{subject to} & \|x\|_2 = 1 \end{array} \tag{2}$$

Problems of this form have a classic solution given by the singular value decomposition (SVD) of A. The solution is the right singular vector corresponding to the smallest singular value of A. If we let σ_i be the i^{th} singular value of A, u_i and v_i be the i_{th} column of U and V respectively, and r be the rank of A

$$A = USV^T = \sum_i^r \sigma_i u_i v_i^T$$

where $U \in \mathbf{R}^{3M\times9}$, $S \in \mathbf{R}^{9\times9}$, and $V \in \mathbf{R}^{9\times9}$. U and V are orthogonal and S is diagonal with its entries the singular values of A. A is assumed to have full rank because we assume noisy measurements. Since an arbitrary vector x in which $\|x\|_2 = 1$ can be decomposed into the columns of U the objective function $\|Ax\|_2^2$ is minimized when x is the same as the v_i corresponding to the smallest σ_i of A.

This solution method is easy to implement, but does not included the information known about J. It is possible with noisy data that the matrix estimated does not satisfy the properties of a valid inertia matrix.

2.2 Projection

One simple method to always estimate a valid inertia matrix is to first use the SVD approach to estimate an inertia matrix, and then project the estimate onto the space of valid inertia matrices. If we let J_{SVD} be the inertia matrix estimated from the SVD approach we can find the projection J_{P} by solving

$$\begin{array}{ll} \text{minimize} & \|J_{\text{P}} - J_{\text{SVD}}\|_F \\ \text{subject to} & J_{\text{P}} \succ 0 \\ & \lambda_1(J_{\text{P}}) \leq \lambda_2(J_{\text{P}}) + \lambda_3(J_{\text{P}}) \end{array} \tag{3}$$

where the variable is J_{P}. This is a convex problem and can be solved easily. Further, it has an analytical form which makes it easy to implement. J_{P} will always be a valid inertia matrix.

2.3 Linearize and Iterate

This method directly includes the constraints on J in the estimation by linearizing and iterating. The idea is to replace the non-convex equality constraint $\|x\|_2 = 1$ present in (1) with the convex constraint $x^T x_k = 1$ where the variable is x and x_k is the solution from the previous iteration. The solution over an iteration becomes x_{k+1}. Iterations are performed until the change in x from one iteration to the next is under a small threshold, or a certain number of maximum iterations is reached. To attempt to make the linearization valid over an iteration a "trust region" constraint is added that $\|x - x_k\| < T$, where T is a small number. If this constraint makes the problem infeasible T is increased until the problem is feasible. When the algorithm is done iterating its objective value is evaluated. This value is an upper bound on the objective value for the original non-convex problem (1) and is compared against a lower bound. If they match then the solution found can be guaranteed to be optimal. Note that this method is only a heuristic. It contains no guarantees on convergence or optimality, but seems to work well in simulation. The iterative method has a simple algorithm.

$$\begin{array}{lll} x_{k+1} = & \text{argmin} & \|Ax\|_2^2 \\ & \text{subject to} & x^T x_k = 1 \\ & & J \succ 0 \\ & & \lambda_1(J) \leq \lambda_2(J) + \lambda_3(J) \\ & & \|x - x_k\| < T \end{array} \tag{4}$$

Obtaining a lower bound

To find a lower bound on (1) a relaxation is formed. The relaxation method used here transforms the original problem using semidefinite programming. This semidefinite programming (SDP) relaxation makes the problem convex by expanding the feasible region. This means that the set of points that are valid for the original problem is a subset of the points that are valid for the relaxation, so the objective value for the relaxation must be less than or equal to the original problem. Starting with (1) an equivalent problem can be formed.

$$\begin{array}{ll} \text{minimize} & x^T A^T A x \\ \text{subject to} & \|x\|_2 = 1 \\ & J \succ 0 \\ & \lambda_1(J) \le \lambda_2(J) + \lambda_3(J) \end{array} \qquad (5)$$

Another equivalent problem can be formed by defining $X = xx^T$, $P = A^T A$, and recalling that $x^T P x = \mathbf{Tr}\, P(xx^T)$. This formulation is equivalent to (1), but now has as variables x and X.

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\,(PX) \\ \text{subject to} & \mathbf{Tr}\,(X) = 1 \\ & X = xx^T \\ & J \succ 0 \\ & \lambda_1(J) \le \lambda_2(J) + \lambda_3(J) \end{array} \qquad (6)$$

At this stage the problem is relaxed by replacing the non-convex equality constraint $X = xx^T$ with the convex positive semi-definite constraint $X - xx^T \succeq 0$. This can be formulated as a Shur Complement [BV04].

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\,(PX) \\ \text{subject to} & \mathbf{Tr}\,(X) = 1 \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0 \\ & J \succ 0 \\ & \lambda_1(J) \le \lambda_2(J) + \lambda_3(J) \end{array} \qquad (7)$$

If one looks closely at (7) the only variable in the objective is X, and the value for x can be made arbitrarily small without violating a constraint. Therefore an equivalent optimization problem can be formed as

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\,(PX) \\ \text{subject to} & \mathbf{Tr}\,(X) = 1 \\ & X \succeq 0 \end{array} \qquad (8)$$

which has the same objective value as (1) without the constraints on J, thus will have the same objective value as the SVD solution. Because of this we expect the lower bound to be tight to the optimal value if the inertia matrix estimated by the SVD approach is valid, and loose otherwise. It is interesting to consider forming the classic Lagrangian dual of (1) to obtain a lower bound instead of forming the SDP relaxation. One may ask what is the relation between the Lagrangian dual and the SDP relaxation? In particular, is one of the bounds better than the other? The answer turns out to be simple: they are duals of each other, and so (assuming a constraint qualification holds) the bounds are exactly the same [Ad03].

3 Results

The solution methods were tested in simulation. A rigid body undergoing torque free motion was simulated from which attitude and attitude rates were obtained. Gaussian noise was added to the attitude rates. After a number of measurements had been made an inertia matrix was estimated using each of the methods discussed. The error for a given estimation was the difference in Frobenius norm from the estimated inertia matrix and the true matrix. Figure 1 displays the results for the estimation error versus number of measurements for each method. The results are interesting in several ways. As expected, once the amount of mea-

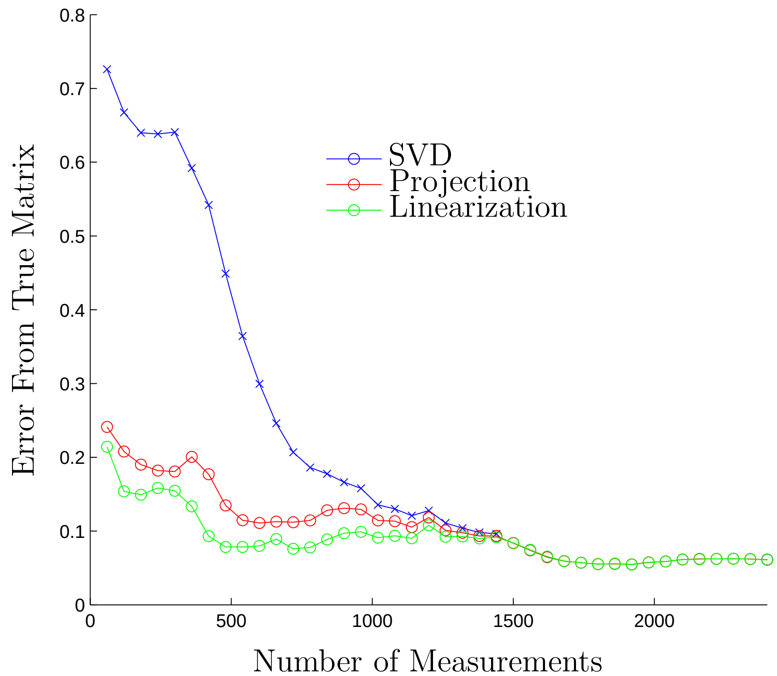

Figure 1: Error from estimated and true inertia matrix versus number of measurements. A ◯ at a data point represents that a physical matrix was estimated, while a × represents a non-physical matrix was estimated.

surements becomes large all methods estimate the same matrix. In this high measurement region the inertia matrix estimated with the SVD approach is a valid one, and the iterative method algorithm exits with a certificate of optimality. When the number of measurements is small the estimates for the inertia matrix are quite different. The SVD approach estimates non-valid matrices and has more error than the projection and linearization method given an estimation. It is also interesting to examine the objective value over iterations for the linearization method. Figure 2 contains the objective value for the linearization method versus iteration number as well as the lower bound in two cases: the top plot is the case in which the inertia matrix estimated by the SVD approach is valid, and the bottom plot is the case in which the SVD approach estimates a non-valid inertia matrix. In the top plot

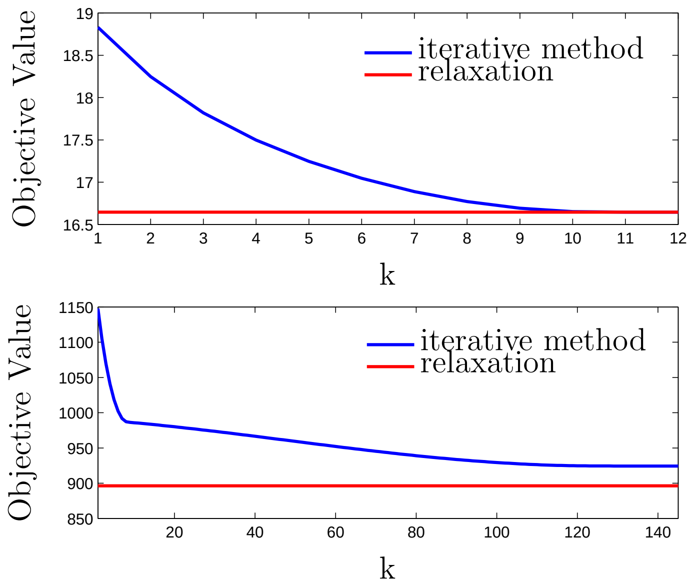

Figure 2: Objective value for linearization method versus iteration number. Also shown is the lower bound from the SDP relaxation. The top plot is a case in which there are relatively high amount of measurements and the inertia matrix estimated by the SVD approach is valid. The bottom plot is the case in which there are fewer measurements and the matrix estimated by the SVD approach is non-valid.

representing the high amount of measurements region the objective value decreases until converging to the lower bound, and thus exits with a certificate of optimality. In the bottom plot representing the low amount of measurements region the objective value converges to a value higher than the lower bound, however, since the lower bound is known to be loose in this region it is possible the linearization method does indeed find the solution. The plots show steady convergence in each case, but recall that the linearization method is only a heuristic.

4 Conclusion

From the simulation data it appears that including known information about the inertia matrix into the estimation produces better estimates in the presence of a low amount of measurements. The projection method provides a quick and easy solution to estimate a valid inertia matrix if the SVD method estimates a non-valid one. The linearization method, while only a heuristic, seems to work well in practice.

References

[Ad03] Stephen Boyd Alexandre d'Aspremont. Relaxations and randomized methods for nonconvex qcqps. 2003.

[BV04] Stephen Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[FL06] A. Feldman and A. Y. Lee. Estimation of state, shape, and inertial parameters of space objects from sequences of range images. *Proceedings of the 16th AAS/AIAA Space Flight Mechanics Conference*, 2006.

[Gre88] D. T. Greenwood. *Principles of Dynamics*. Prentice-Hall, 1988.

[SR09] Daniel Sheinfeld and Steve Rock. Rigid body inertia estimation with applications to the capture of a tumbling satellite. *Proceedings of 19th AAS/AIAA Spaceflight Mechanics Meeting*, pages 343–356, 2009.

[Zam96] M. Zampato. Visual motion estimation for tumbling satellite capture. *Proceedings of The British Machine Vision Conference*, 1996.