System Identification of DragonFly UAV via Bayesian Estimation

Sun Hwan Lee
Department of Aeronautics and Astronautics
Stanford University
Stanford, CA 94305
sunhlee@stanford.edu

Youn Mi Park
Department of Aeronautics and Astronautics
Stanford University
Stanford, CA 94305
younmi@stanford.eud

Abstract

The system identification of UAV, which is the estimation of the parameters of the equation of motion, is crucial for implementing the autopilot. The system identification especially with noisy data is challenging problem. In this report, we will estimate parameters of equation of motion for DragonFly UAV from Stanford Hybrid Systems Laboratory in Aero/Astro department via Baysian estimation. The basic idea here is that the most probable coefficients of the equation of motion, which is set to the unknown set, can be found by maximizing the probability of that unknown set given the data. Furthermore, from Bayesian rule, the conditional probability is proportional to the product of the likelihood function and the prior. By properly defining the likelihood function and the prior, we will be able to find the most probable unknown set of parameters. Identifying the system of UAV, the performance of the estimator will be evaluated by cross validation and in particularly the benefit of the Bayesian estimation will be highlighted by comparing it to the general least square method.

1 Bayesian System Identification

The system identification is the first and crucial step for the design of the controller, simulation of the system and so on. Frequently it is necessary to analyze the flight data in the frequency domain to identify the UAV sytem. Our approach, in this project, uses the data from the flight test of the UAV in time domain thus does not require to excite every mode of a given system. This approach is beneficial in terms of the simpleness and computational efficiency as Bayesian estimator can be performed with less computational cost than frequency domain least square method. In addition, the theoretical derivation is not limited to the linear system thus it can be easily extended to the nonlinear system identification when the assumption imposed on the following theoretical background works well.

Suppose the system, which is linear or nonlinear, is given.

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x};\mathbf{c}) + \mathbf{u}(t) + \omega(t) \quad (1)$$

$$\mathbf{y}(t) = \mathbf{x}(t) + \nu(t). \quad (2)$$

where $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^p$, $\mathbf{c} \in \mathbb{R}^k$ are state variable, system output and unknown coefficient vector to be determined. Also the dynamic noise and measurement noise are assumed as white Gaussian noise which satisfies the following property.

$$E\left[\omega(t)\right] = 0, \quad \text{cov}\left(\omega(t)\right) = \hat{D}, \quad (3)$$

$$E\left[\nu(t)\right] = 0, \quad \text{cov}\left(\nu(t)\right) = \sigma^2 \hat{I}. \quad (4)$$

where ω, $\nu \in \mathbb{R}^p$.

Another assumption we will use through this report is that the full states were measured, thus $\omega(t)$ can represent both the dynamic noise and the measurement noise. From this assumption our unknown set is reduced to $\mathcal{M} = \{\mathbf{c},\, \hat{D}\}$. The purpose of this paper, which is to find out the most probable unknown set given the time series of data, can be achieved by choosing the unknown set that gives the peak value of posterior probability, $P_{ps}(\mathcal{M}|\mathcal{Y})$. From Bayes' theorem,

$$P_{ps}(\mathcal{M}|\mathcal{Y}) = \frac{P\left(\mathcal{Y}|\mathcal{M}\right) P_{pr}\left(\mathcal{M}\right)}{\int P\left(\mathcal{Y}|\mathcal{M}\right) P_{pr}\left(\mathcal{M}\right) d\mathcal{M}}. \tag{5}$$

The basic idea is to update the posterior of the unknown set by using the time sequence of data and replace the prior by the posterior probability distribution. The theoretical detail and the application in the case of the stochastic system driven by only white Gaussian noise is well provided in [1]. But the application of Bayesian inference to the practical system identification is meaningful task. The noise property has a key role to construct the probability density function for the state variable, that is time series of data, in our report.

1.1 Maximum Likelihood Estimation

The midpoint Euler discretization scheme is used to construct the discrete system.

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{f}(\tilde{\mathbf{x}}_n; \mathbf{c}) + h\mathbf{u}_n + \mathbf{z}_n \tag{6}$$

$$\mathbf{y}_n = \mathbf{x}_n. \tag{7}$$

where $E\left[\mathbf{z}_n,\, \mathbf{z}_{n'}^T\right] = h^2 \hat{D}\delta_{nn'}$. The last equality came from the assumption about $\omega(t)$ which is previously mentioned. From this point, the state variable, $\mathbf{x}_n$, and the system output, $\mathbf{y}_n$, are set to be identical. The procedure toward the most probable unknown set is as follows. The likelihood function, $P\left(\mathcal{Y}|\mathcal{M}\right)$ is obtained by transformation of the probability density function of $\mathbf{z}_n$ to the function of the output, $\mathbf{y}_n$, along the given data. Adding initial guess for the prior distribution of the unknown set gives the posterior distribution of the unknown set which has to be maximized.

The probability density function of $\mathbf{z}_n$ along the given data is

$$P\left[\{\mathbf{z}_n\}\right] = \prod_{n=0}^{m-1} \frac{1}{\sqrt{(2\pi)^p \left|h^2 \hat{D}\right|}} \exp\left(\frac{1}{2h^2}\mathbf{z}_n^T \hat{D}^{-1}\mathbf{z}_n\right). \tag{8}$$

The transformation of the probability function from $\mathbf{z}_n$ to $\mathbf{x}_n$ is related by Jacobian matrix.

$$P\left[\{\mathbf{x}_n\}\right] = \frac{d\mathbf{z}_n}{d\mathbf{x}_{n+1}} P\left[\{\mathbf{z}_n\} \rightarrow \{\mathbf{x}_n\}\right]. \tag{9}$$

Jacobian matrix is

$$\frac{d\mathbf{z}_n}{d\mathbf{x}_{n+1}} = \prod_{n=1}^{m}\prod_{i=1}^{p}\left[1 - \frac{h}{2}\frac{\partial f_i\left(\tilde{\mathbf{x}}_{n-1}; \mathbf{c}\right)}{\partial \mathbf{x}_i^n}\right] \approx \exp\left[-\frac{h}{2}\sum_{n=1}^{m} \mathrm{tr}\Phi\left(\tilde{\mathbf{x}}_{n-1}; \mathbf{c}\right)\right]. \tag{10}$$

where, $\tilde{\mathbf{x}}_j = \frac{\mathbf{x}_{j+1}+\mathbf{x}_j}{2}$ and $\Phi_{ij}\left(\mathbf{x}; \mathbf{c}\right) = \frac{\partial f_i(\mathbf{x};\mathbf{c})}{\partial \mathbf{x}_j}$. This assumption here is crucial to make the objective function be a convex problem even in the case of the nonlinear system. Thus a relatively simple system such as Lorenz system can be identified quite accurately [1].

Substituting Eq.(8) and (10) into Eq.(9) and the probability of the initial state produces the probability density function of the state variable $\mathbf{y}_n$ along the given data.

$$\begin{aligned} P\left(\mathcal{Y}|\mathcal{M}\right) &= P\left(\mathbf{y}_0\right) \exp\left[-\frac{h}{2}\sum_{n=0}^{m-1} \mathrm{tr}\Phi\left(\tilde{\mathbf{y}}_n; \mathbf{c}\right)\right] \prod_{n=0}^{m-1} \frac{1}{(2\pi)^{p/2}\left|h^2 \hat{D}\right|^{1/2}} \\ &\quad \exp\left(-\frac{1}{2}\left[\dot{\mathbf{y}}_n - \hat{f}\left(\tilde{\mathbf{y}}_n; \hat{\mathbf{c}}\right)\right]^T \hat{D}^{-1}\left[\dot{\mathbf{y}}_n - \hat{f}\left(\tilde{\mathbf{y}}_n; \hat{\mathbf{c}}\right)\right]\right). \end{aligned} \tag{11}$$

where, $\dot{\mathbf{y}}_n = \frac{\mathbf{y}_{n+1}-\mathbf{y}_n}{h}$. Now, $\hat{f}$ and $\hat{c}$ is the equation and unknown coefficients including input signal and input coefficients.

With the initial guess of Gaussian in $\hat{\mathbf{c}}$ and uniform distribution in $\hat{D}$, the posterior of the unknown set, which is the objective function is explicitly expressed.

$$\ell(\hat{\mathbf{c}}, \hat{D}) = \log L(\hat{\mathbf{c}}, \hat{D}) = \log P\left(\mathcal{M}|\mathcal{Y}\right) \propto -\rho(\hat{D}) + \hat{\mathbf{c}}^T\varphi(\hat{D}) - \frac{1}{2}\hat{\mathbf{c}}^T\Lambda(\hat{D})\hat{\mathbf{c}}, \tag{12}$$

$$\rho(\hat{D}) = \frac{m}{2}\log|\hat{D}| + \frac{1}{2}\sum_{n=0}^{m-1}\dot{\mathbf{y}}_n^T\hat{D}^{-1}\dot{\mathbf{y}}_n, \tag{13}$$

$$\varphi(\hat{D}) = \Sigma_{pr}\hat{\mathbf{c}}_{pr} + \sum_{n=0}^{m-1}\left(U_n^T\hat{D}^{-1}\dot{\mathbf{y}}_n - \frac{h}{2}\mathcal{U}_n^T\right), \quad \Lambda(\hat{D}) = \Sigma_{pr} + \sum_{n=0}^{m-1}U_n^T\hat{D}^{-1}U_n. \tag{14}$$

where, $\hat{f}(\mathbf{x};\mathbf{c}) = U(\mathbf{x})\hat{\mathbf{c}}, \quad \mathcal{U}_n = \left[\frac{\partial\hat{f}_1(\tilde{\mathbf{x}}_n)}{\partial\mathbf{x}_1} \quad \ldots \quad \frac{\partial\hat{f}_p(\tilde{\mathbf{x}}_n)}{\partial\mathbf{x}_p}\right]$.

The unknown set, $\mathcal{M}$, maximizing the objective function can be obtained by iterating the maximization until the convergence. For the first iteration, $\hat{\mathbf{c}}_{pr}$ is used as the first guess for $\hat{\mathbf{c}}$ and the unknown set is updated subsequently while Σ_{pr} is updated as $\Lambda(\hat{D})$. The maximum likelihood estimate of $\hat{D}$ and $\hat{c}$ are

$$\hat{D} = \frac{1}{m}\sum_{n=0}^{m-1}\left(\dot{\mathbf{y}}_n - U_n\hat{\mathbf{c}}\right)\left(\dot{\mathbf{y}}_n - U_n\hat{\mathbf{c}}\right)^T, \quad \hat{\mathbf{c}} = \Lambda(\hat{D})^{-1}\varphi(\hat{D}). \tag{15}$$

2 Equation of Motion of DragonFly

The linearized equation of motion of DragonFly [2] is the hypothesis of the learning algorithm.

$$\dot{V}_T = A_{Xw} \tag{16}$$

$$\dot{\alpha} = q - (p\cos\alpha + r\sin\alpha)\tan\beta + \frac{A_{Zw}}{V_T\cos\beta} \tag{17}$$

$$\dot{\beta} = -(r\cos\alpha - p\sin\alpha) + \frac{A_{Yw}}{V_T} \tag{18}$$

$$\dot{p} = \frac{I_y - I_z}{I_x}qr + \frac{\tilde{q}Sb}{I_x}C_l \tag{19}$$

$$\dot{q} = \frac{I_z - I_x}{I_y}rp + \frac{\tilde{q}Sc}{I_y}C_m \tag{20}$$

$$\dot{r} = \frac{I_x - I_y}{I_z}pq + \frac{\tilde{q}Sb}{I_z}C_n \tag{21}$$

$$\begin{bmatrix}\dot{\phi}\\ \dot{\theta}\\ \dot{\psi}\end{bmatrix} = \begin{bmatrix}1 & \sin\phi\tan\theta & \cos\phi\tan\theta\\ 0 & \cos\phi & -\sin\phi\\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta\end{bmatrix}\begin{bmatrix}p\\ q\\ r\end{bmatrix} \tag{22}$$

Eq.(19) - (21) are referred to the wind axis coordinate and Eq.(22) - (24) are established about the body fixed coordinate. $(V_T,\, \alpha,\, \beta)$ are the total velocity, angle of attack and the sideslip angle, $(p,\, q,\, r)$ are angular velocity in the body fixed coordinate, $(\phi,\, \theta,\, \psi)$ are Euler angle. Also $(A_{Xw},\, A_{Yw},\, A_{Zw})$ represents the force in the wind axis and $\tilde{q}$, S, b, c are dynamic pressure, wing platform area, wing span and mean chord. Because the data from the flight data is based on the small perturbation deviated from the nominal condition, the hypothesis of the data turns out to be the linearized equation of motion, Eq.(23).

$$\begin{bmatrix}\Delta\dot{v}_T\\ \Delta\dot{\alpha}\\ \dot{q}\\ \Delta\dot{\theta}\\ \Delta\dot{\beta}\\ \dot{p}\\ \dot{r}\\ \Delta\dot{\phi}\end{bmatrix} = A\begin{bmatrix}\Delta v_T\\ \Delta\alpha\\ q\\ \Delta\theta\\ \Delta\beta\\ p\\ r\\ \Delta\phi\end{bmatrix} + B\begin{bmatrix}\delta_e\\ \delta_t\\ \delta_a\\ \delta_r\end{bmatrix} \tag{23}$$

where, A and B matrix are the coefficients matrix which will be estimated and δ_e, δ_t δ_a δ_r are elevator input, throttle input, aileron input and rudder input.

3 System Identification of DragonFly

In this section, the result of dynamics system identification will be presented by comparing the flight data and trajectory from the estimated model. The result of the system identification is presented in Fig.(1) and (2). These figures show the result of longitudinal and lateral system identification.

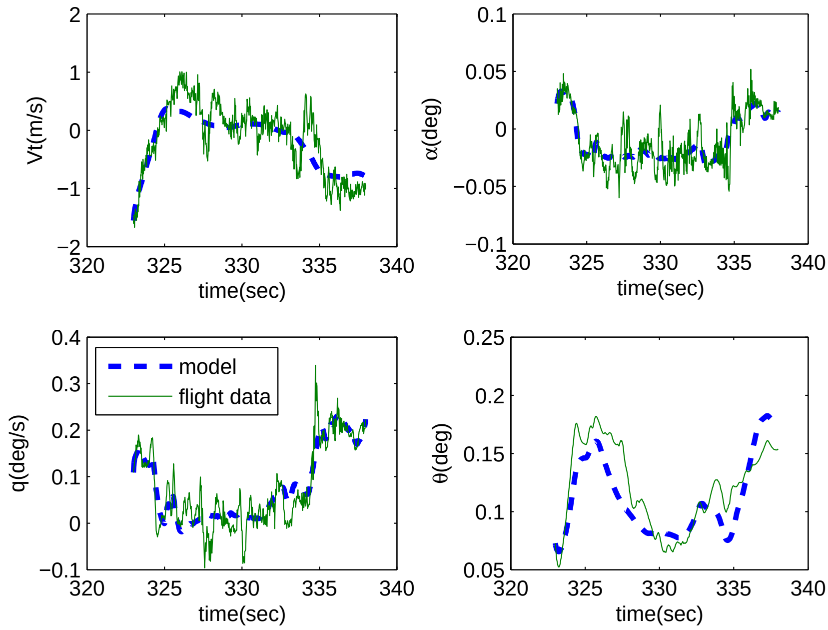

Figure 1: Longitudinal system dynamics identification

4 Cross Validation and Comparison with GLS

The cross validation of Bayesian system identification was performed by utilizing different flight data from different dates and computing the averaged sum of squared error. To emphasize the benefit of Bayesian approach the sum of squared error of each state variable was compared with one from GLS(Generalized Least Square) method. Table (1) represents the comparison of sum of squared error between Bayesian and GLS method.

Table 1: The averaged sum of squared error of Bayesian and GLS.

State variable	GLS	Bayesian
v_T	6.4690	2.4328
α	0.0009	0.0002
q	0.0093	0.0020
θ	0.0909	0.0039
β	0.0023	0.0003
p	0.0207	0.0127
r	0.0813	0.0070
ϕ	0.2531	0.0591

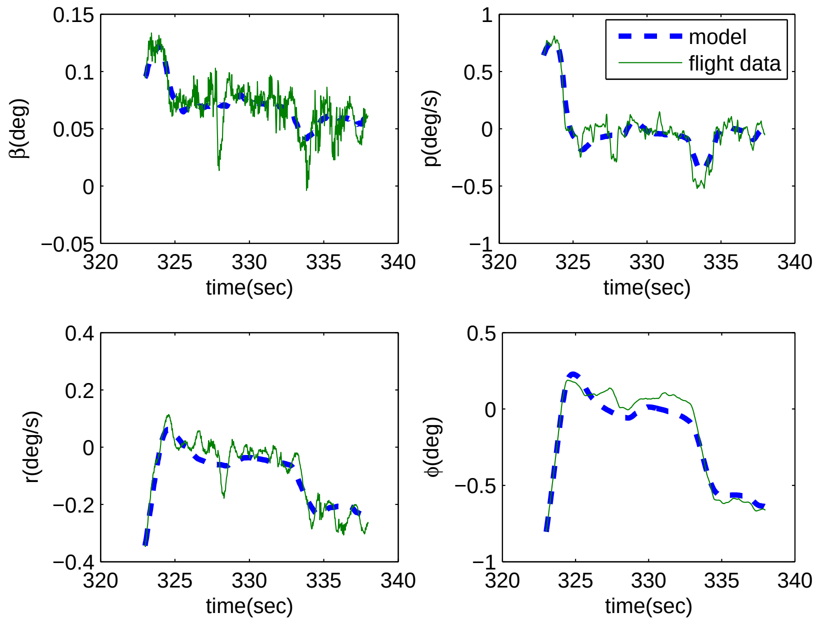

Figure 2: Lateral system dynamics identification

5 Conclusion and Future Work

Combined with the simulated trajectory using the estimated model and comparison of sum of squared error the conclusion can be drawn that the estimated system using Bayesian approach tracks the measured data quite well and beats least square method in time domain in terms of the averaged sum of squared error. In addition, the pure system property can be obtained as Bayesian system identification can extract the information about the measurement noise. In other words, Bayesian approach is more appropriate when it comes to the system identification from very noisy measurement.

Other advantage of this approach is that Bayesian approach can be extended to nonlinear systems. Although the extension to the nonlinear system equation was limited because some of coefficients depend on state variables, Bayesian approach can be applied to the identification of nonlinear systems whose coefficients are constant.

It is well understood that the best way to test the performance of the system identification is to implement the controller based on the identified model. This method was impossible due to various reasons such as the limitation of time and cost. We would like to say that, however, the controller execution to the real dynamic system is interesting and valuable future work.

References

[1] Smelyanskiy, V.N. & Luchinsky, D. G. & Timucin, D. A. & Bandrivskyy, A. (2005) Reconstruction of stochastic nonlinear dynamical models from trajectory measurements. *Physical Review E* **72**(026202):1-12.

[2] Jang, Jung Soon and Tomlin, Claire J. (2001) Autopilot Design for the Stanford DragonFly UAV: Validation through Hardware-in-the-Loop Simulation. In *AIAA Guidance, Navigation, and Control Conference*, Montreal, QB, Canada, August.