Robust filtering with uncertainty in initial state

Nithin Sugavanam and Emre Ertin
Department of Electrical and Computer Engineering
The Ohio State University

Problem setup

- We consider estimation of $X_0 \sim \sum_{i=1}^{K} \alpha_i \mathcal{N}(\mu_i, \Sigma_i)$ using a linear observation model $Z_0 = H_0 X_0 + N_0$,for example: target initial state with Gaussian noise.
- We assume that the prior distribution α on the components is not known.
- Goal is to find a class of estimators minimizing the mean square error that are robust to uncertainty in prior distribution.
- We also consider the problem of observation matrix design for a class of estimators robust to uncertainty in α that minimize the mean square error.

Sequential estimation / filtering with mixture of Gaussians prior

- We consider the following Gauss-Markov model for the target $X_{t+1} = F_t X_t + W_t$ with the observation $Z_{t+1} = H_t X_t + N_t$.
- The mean square error for a given α and an estimator $\hat{x}$ at some time T is

$$M(\alpha, \hat{x}) = \sum_{t=0}^{T} E\left(\|x_t - \hat{x}_t\|^2\right).$$

- For the minimum mean square error criterion, the optimal non-linear estimator for a given α takes the form of a mixture of Kalman filters (for each Gaussian component), which is combined using a non-linear function of the observation.
- The update equations for non-linear filter can be written as

$$\alpha^{(i)}_{(t|t)} = \frac{\alpha^{(i)}_{(t|t-1)} \mathcal{N}_{z_t}\left(H_t \mu^{(i)}_{(t|t-1)} + \mu_{N_t}, H_t \Sigma^{(i)}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right)}{\sum_k \alpha^{k}_{(t|t-1)} \mathcal{N}_{z_t}\left(H_t \mu^{(k)}_{(t|t-1)} + \mu_{N_t}, H_t \Sigma^{(k)}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right)}, \quad \alpha^{(i)}_{(t+1|t)} = \alpha^{(i)}_{(t|t)}$$

$$\mu^{(i)}_{(t|t)} = \mu^{(i)}_{(t|t-1)} + \Sigma^{(i)}_{(t|t-1)} H_t^T \left(H_t \Sigma^{(i)}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right)^{-1} \left(z_t - H_t \mu^{(i)}_{(t|t-1)} - \mu_{N_t}\right), \quad \mu^{(i)}_{(t+1|t)} = F_{t+1} \mu^{(i)}_{(t|t)} + \mu_{w_t}$$

$$\Sigma^{(i)}_{(t|t)} = \Sigma^{(i)}_{(t|t-1)} - \Sigma^{(i)}_{(t|t-1)} H_t^T \left(H_t \Sigma^{(i)}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right)^{-1} H_t \Sigma^{(i)}_{(t|t-1)}, \quad \Sigma^{(i)}_{(t+1|t)} = H_{t+1} \Sigma^{(i)}_{(t|t)} H_{t+1}^T + \Sigma_{W_{t+1}}$$

$\hat{x}_{(t|t)} = \sum_i \alpha^{(i)}_{(t|t)} \mu^{(i)}_{(t|t)}$ is the minimum mean square error estimator.

- When we restrict it to a family of linear estimators, the optimal LMMSE estimator is the Kalman filter for a Gaussian distribution for initial state, which has the same mean and covariance as the mixture of Gaussians.

$$\bar{\mu}_0 = \sum_i \alpha_i \mu_i \quad \bar{\Sigma}_0 = \sum_i \alpha_i \left[\Sigma_i + \mu_i \mu_i^T\right] - \bar{\mu}_0 \bar{\mu}_0^T$$

$$\bar{x}_{(t|t)} = \bar{x}_{(t|t-1)} + \bar{\Sigma}_{(t|t-1)} H_t^T \left[H_t \bar{\Sigma}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right]^{-1} \left(z_t - H_t \bar{x}_{(t|t-1)} - \mu_{N_t}\right)$$

- The mean square error for the linear estimator given α at time T can be written as

$$M^*(\alpha) = \sum_{t=0}^{T} Tr\left[\bar{\Sigma}_{(t|t-1)} - \bar{\Sigma}_{(t|t-1)} H_t^T \left[H_t \bar{\Sigma}_{(t|t-1)} H_t^T + \Sigma_{N_t}\right]^{-1} H_t \bar{\Sigma}_{(t|t-1)}\right]$$

Robust filtering problem

- The minimax robust filter is given by

$$\bar{x}^* = \arg \min_{\bar{x} \in C_L} \max_{\alpha \in P} M(\alpha, \bar{x})$$

- The least favorable prior distribution is given by

$$\bar{\alpha}^* = \arg \max_{\alpha \in P} \min_{\bar{x} \in C_L} M(\alpha, \bar{x})$$

- The filter and prior distribution $(\alpha_L, \bar{x}_L)$ form a saddle point iff

$$M(\alpha, \bar{x}_L) \le M(\alpha_L, \bar{x}_L) \le M(\alpha_L, \bar{x}).$$

- If α^* forms a saddle point with linear filter $\bar{x}(\alpha^*)$, then $\bar{x}(\alpha^*)$ belongs to the set of robust minimax filters.
- The prior distribution and filter $(\alpha_L, \bar{x}_L)$ form a regular pair iff for any $\alpha \in B_\epsilon(\alpha_L)$ $M^*(\alpha) - M(\alpha_L, \bar{x}(\alpha)) = o(\epsilon)$.

Existence of saddle point for linear filters

- Let P be the probability simplex, C_L be the space of linear filters.
- It can be verified that $(\alpha_L, \bar{x}_L)$ form a regular pair.
- $M(., \bar{x})\ \forall \bar{x} \in C_L$ is concave in α.
- P is a convex set. Hence by invoking Theorem in [1], it can be concluded that $(\alpha_L, \bar{x}_L)$ form a saddle point solution if α_L exists.
- It can be seen that for any class of filters $V(\alpha) = \min_{\hat{x} \in C} E\left(\|x - \hat{x}\|^2\right)$ is convex α.

Robust linear observation matrix design

- Given model $Z_0 = H_0 X_0 + N_0$, goal is to obtain H_0 with Frobenius norm constraints , which is robust to uncertainty in α such that,

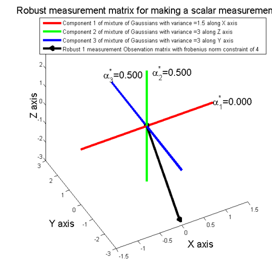
$$H_0^* = \arg \max_{H_0 \in H} \min_{\hat{x} \in C} \max_{\alpha \in P} E\left(\|X_0 - \hat{x}(Z_0)\|^2\right).$$

- Since for the linear class of filters, the least favorable mixing prior, and the linear filter form a saddle point, we can interchange the order to obtain

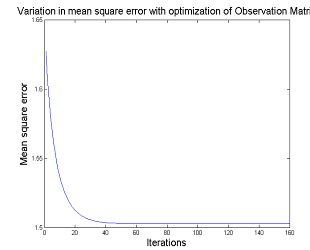
$$H_0^* = \arg \min_{H_0 \in H} \max_{\alpha \in P} Tr\left[\bar{\Sigma}_0(\alpha) - \bar{\Sigma}_0(\alpha) H_0^T \left[H_0 \bar{\Sigma}_0(\alpha) H_0^T + \Sigma_{W_0}\right]^{-1} H_0 \bar{\Sigma}_0(\alpha)\right].$$

- The inner optimization problem is solved for each H_0 .
- Using Envelope theorem, we obtain the derivative of $V(H_0) = \max_{\alpha \in P} Tr\left[\bar{\Sigma}_0(\alpha) - \bar{\Sigma}_0(\alpha) H_0^T \left[H_0 \bar{\Sigma}_0(\alpha) H_0^T + \Sigma_{W_0}\right]^{-1} H_0 \bar{\Sigma}_0(\alpha)\right]$ with H_0.
- By utilizing projection methods and gradient with H_0, we optimize the linear observation matrix.

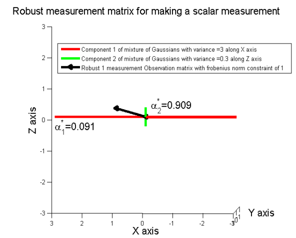
Examples



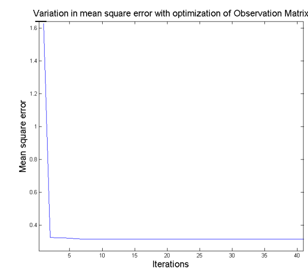
1 dimensional observation model with three mixture components of unequal variance.



Mean square error variation for the case of 3 mixture components with unequal variance



1 dimensional observation model with two mixture components of unequal variance.



Mean square error variation for the case of 2 mixture components with unequal variance

Future Work

- Investigate the existence of saddle point for minimum mean square error estimator.
- Observation matrix design using information theoretic framework using variational approximations.

References

- [1] "On minimax robustness: A general approach and applications.", S. Verdu, H.V. Poor, IEEE Transactions on Information Theory 30(2):328-340 (1984)
- [2] " Knowledge enhanced compressie measurement designs for estimating sparse signals in clutter", S. Jain. A. Soni, J. Haapt, N. Rao, R. Nowak, in Proc. Signal Processing with Adaptive Sparse Structured presentations (SPARS), 2013, to appear.