Sensors networks

Estimation

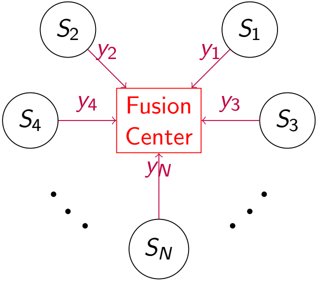
Manuel A. Vázquez
Joaquín Míguez
Jose Miguel Leiva

February 4, 2024

Index

Index

Centralized estimation network



Structure of the network with N sensors ($i = 1, \ldots, N$):

- $S_i \equiv i$-th sensor
- $y_i \equiv$ observation coming from the i-th sensor

The FC must use the collection of observations to estimate an $M \times 1$ vector, $\mathbf{x}$.

Classic estimation

Estimation of $\mathbf{x}$ given the collection of data $\mathbf{y} = \{y_1, ..., y_N\}$ is tantamount to a *classic* estimation problem. There are several possible estimators:

- Maximum likelihood (ML)

$$\hat{\mathbf{x}}^{ML} = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}).$$

- Maximum a posteriori (MAP)

$$\hat{\mathbf{x}}^{MAP} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}).$$

- Minimum Mean Square Error (MMSE)

$$\begin{aligned} \hat{\mathbf{x}}^{MMSE} &= \arg\min_{\hat{\mathbf{x}}} \mathbb{E}\left[\|\mathbf{x} - \hat{\mathbf{x}}\|^2\right] \\ &= \mathbb{E}\left[\mathbf{x}|\mathbf{y}\right] = \int \mathbf{x} p(\mathbf{x}|\mathbf{y}) d\mathbf{x}. \end{aligned}$$

Index

Dynamic model

A more interesting problem...

What if the variable of interest changes with time?

$$\mathbf{x} \rightarrow \mathbf{x}_t$$

with t being a discrete-time index. Since $\mathbf{x}$ was a random variable, $\mathbf{x}_t$ is a *random process*.

Goal

We want to **track** the evolution of $\mathbf{x}$ with time.

Then, we need two equations

- a state equation modeling the evolution of the variable of interest
- a observation equation modeling the connection between the variable of interest and the observations

Let us start by considering an easy-to-handle model...

Linear Gaussian model

- Process $\mathbf{x}_t$ evolves according to a linear Gaussian model

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{v}_t \qquad \text{(state equation)}$$

where $\mathbf{F}$ is a $M \times M$ matrix, and $\mathbf{v}_t$ is a $M \times 1$ Gaussian random vector with mean $\mathbf{0}$ and covariance $\mathbf{Q}$, being M the number of elements in $\mathbf{x}_t$.

- Connection between the variable of interest $\mathbf{x}_t$ and the observations is given by

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}_t + \mathbf{w}_t \qquad \text{(observation equation)}$$

where $\mathbf{H}$ is a $N \times M$ matrix and $\mathbf{w}_t$ is a $N \times 1$ Gaussian random vector with mean $\mathbf{0}$ and covariance $\mathbf{R}$.

Observations...

...vector, $\mathbf{y}_t$, collects the measurements from all the sensors

Example I

Goal

Localization and tracking of an object moving with a *known* constant velocity $\mathbf{c}$.

- The position of the target, $\mathbf{x}_t$ (here representing the state of the system), evolves according to

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{c}T + \mathbf{v}_t,$$

where T is the *sampling period*, i.e., the time elapsed between two consecutive observations.

- The position is observed directly:

$$\mathbf{y}_t = \mathbf{x}_t + \mathbf{w}_t$$

Example II

Goal

Localization and tracking of an object moving with constant *unknown* velocity.

In order to estimate the velocity, it is included in the state of the system

$$\mathbf{x}'_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{c} \end{bmatrix}$$

- The state equation is now

$$\mathbf{x}'_t = \mathbf{F}\mathbf{x}'_{t-1} + \begin{bmatrix} \mathbf{v}_t \\ \mathbf{0} \end{bmatrix}, \text{ with } \mathbf{F} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

- and the observation equation

$$\mathbf{y}_t = \mathbf{H}\mathbf{x}'_t + \mathbf{w}_t, \text{ with } \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Index

What is the Kalman filter?

It is a recursive method for computing *posterior* probability distributions in linear Gaussian dynamic systems.

- recursive: at every time instant it yields an estimate based on the one from the previous time instant
- probability distributions: it does not (only) give us an estimate of the parameter of interest, but rather its full distribution (determined by the mean and the covariance matrix)
- dynamic systems: systems that model a time-varying magnitude
- linear: the equations defining the system are linear with respect to the variable of interest
- Gaussian: the noise affecting the above equations is Gaussian.

Dynamic system in state-space form

$$\left.\begin{array}{lcc} \mathbf{x}_t = \mathbf{F}_{t-1}\mathbf{x}_{t-1} + \mathbf{v}_t & \longleftarrow & \text{state equation} \\ \mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t & \longleftarrow & \text{observation equation} \end{array}\right\} \text{dynamic system in state-space form}$$

Goal

Recursively estimate the state of the system $\mathbf{x}_t$ given the observations, $\mathbf{y}_t$

The time index is discrete, i.e, $t = 0, 1, \cdots$

Notation

$$\mathbf{x}_t = \mathbf{F}_{t-1}\mathbf{x}_{t-1} + \mathbf{v}_t$$

- $\mathbf{x}_t$ is the (system) state vector at time t
- $\mathbf{F}_t$ is the state transition matrix: it determines the evolution of the state of the system
- $\mathbf{v}_t \sim \mathcal{N}(\mathbf{v}_t|\mathbf{0}, \mathbf{Q}_t)$ is the state (or process) noise
  - $\mathbf{Q}_t$ is the covariance matrix for the state noise (it might be time-varying)

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t$$

- $\mathbf{y}_t$ is the observations vector
- $\mathbf{H}_t$ is the observation matrix: it connects the observations with the (unknown) state
- $\mathbf{w}_t \sim \mathcal{N}(\mathbf{w}_t|\mathbf{0}, \mathbf{R}_t)$ is the observation noise
  - $\mathbf{R}_t$ is the covariance matrix of the observation noise (it might be time-varying)

Filtering: initial hypothesis

The *prior* (initial) distribution of the state is Gaussian, i.e.,

$$p(\mathbf{x}_0) \sim \mathcal{N}\left(\mathbf{x}_0 | \hat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}\right)$$

with **known**

- $\hat{\mathbf{x}}_{0|0}$: the mean for the distribution of the state at time 0
- $\mathbf{P}_{0|0}$: covariance matrix for the distr. of the state at time 0

Notation

- $\hat{\mathbf{x}}_{a|b} \equiv$ mean *estimated* at time a using the observations up to time b
- $\mathbf{P}_{a|b} \equiv$ covariance matrix *estimated* at time a using the observations up to time b

In the beginning

At time 0 we have no observations available...but the notation is still convenient.

Filtering: recursion

If the initial hypothesis holds then, for $t = 1, 2, \cdots$

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t-1}) \sim \mathcal{N}\left(\mathbf{x}_t | \hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1}\right) \quad \longleftarrow \quad \text{predictive pdf}$$

$$p(\mathbf{x}_t \mid \mathbf{y}_{1:t}) \sim \mathcal{N}\left(\mathbf{x}_t | \hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t}\right) \quad \longleftarrow \quad \text{filtered pdf}$$

...that is, both pdf's are Gaussian if the initial distribution is Gaussian, and the **Kalman filter** yields both its means and covariances in a two-steps process

Notation

$y_{i:j} \equiv \{y_i, y_{i+1}, \cdots, y_j\}$

Solution

Predictive step

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_{t-1}\hat{\mathbf{x}}_{t-1|t-1} \quad \longleftarrow \quad \text{predictive mean}$$

$$\mathbf{P}_{t|t-1} = \mathbf{Q}_{t-1} + \mathbf{F}_{t-1}\mathbf{P}_{t-1|t-1}\mathbf{F}_{t-1}^{H} \quad \longleftarrow \quad \text{predictive covariance}$$

associated with the predictive pdf

Update step

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^H\left(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^H + \mathbf{R}_t\right)^{-1} \quad \longleftarrow \quad \text{Kalman gain}$$

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}\right) \quad \longleftarrow \quad \text{filtered mean}$$

$$\mathbf{P}_{t|t} = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t\right)\mathbf{P}_{t|t-1} \quad \longleftarrow \quad \text{filtered covariance}$$

associated with the filtered pdf

Remarks

- In order to apply the KF
  - the **system** must be **linear**
  - the **noise** must be **Gaussian**
  - the **initial distribution** of the state must be **Gaussian**
- KF outputs the a probability distribution[1], which contains all the available information about the parameter of interest

In our case...

$p(\mathbf{x}_t \mid \mathbf{y}_{1:t})$ contains all the available information at time t about $\mathbf{x}_t$, and from it we can compute the mean, median, mode...

- $\hat{\mathbf{x}}_{t|t} \equiv$ MMSE estimate of the state at time t
- $\mathrm{Tr}\left\{\mathbf{P}_{t|t}\right\} \equiv$ minimum error (square error at $\hat{\mathbf{x}}_{t|t}$)

[1] ...since a Gaussian distribution is fully specified by its means vector and covariance matrix

Index

State and observation equations

$$\mathbf{x}_t = \mathbf{F}_{t-1}\mathbf{x}_{t-1} + \mathbf{B}_t\mathbf{u}_t + \mathbf{v}_{t-1} \quad \longleftarrow \quad \text{state equation}$$

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{w}_t \quad \longleftarrow \quad \text{observation equation}$$

The control term, $\mathbf{B}_t\mathbf{u}_t$, with

- $\mathbf{B}_t$ is the control-input, and
- $\mathbf{u}_t$ is the control vector

is **known** (at every time instant t) and meant for modifying the (unknown) state.

What is the control term good for?

- ...sometimes we can have some impact (control) over whatever we aim at estimating

example

Problem: estimating the trajectory of a drone we are handling ourselves

- ...from a mathematical standpoint, it is useful to model *affine* functions

The control term is something that affects the state[2], but since it is known there is no need to estimate it.

[2]...so it should be in the corresponding equation!!

Solution

Predictive step

$$\begin{rcases} \hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_{t-1}\hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t\mathbf{u}_t & \longleftarrow \text{predictive mean} \\ \mathbf{P}_{t|t-1} = \mathbf{Q}_{t-1} + \mathbf{F}_{t-1}\mathbf{P}_{t-1|t-1}\mathbf{F}_{t-1}^H & \longleftarrow \text{predictive covariance} \end{rcases} \text{associated with the predictive pdf}$$

Update step

$$\begin{rcases} \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^H\left(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^H + \mathbf{R}_t\right)^{-1} & \longleftarrow \text{Kalman gain} \\ \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}\right) & \longleftarrow \text{filtered mean} \\ \mathbf{P}_{t|t} = \left(\mathbf{I} - \mathbf{K}_t\mathbf{H}_t\right)\mathbf{P}_{t|t-1} & \longleftarrow \text{filtered covariance} \end{rcases} \text{associated with the filtered pdf}$$