Lecture 3: Workability and quality estimation of DNN non-parametric modeling

Plan of presentation

- Attractive ellipsoid (AE) method
- Energetic ellipsoidal function
- State estimation error by DNNO: zone-convergence analysis
- Learning law designing
- Relation of AE with DNNO parameters
- Feedback optimization

Attractive ellipsoid (AE) method

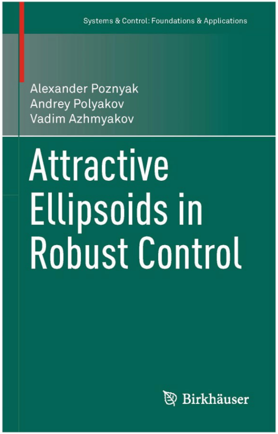

Figure 1: Book on AEM published in 2014.

Lemma

If some differentiable function V_t satisfies the **differential inequality**

$$\boxed{\dot{V}_t \leq -\alpha V_t + \beta, \alpha > 0, t \geq 0,}$$

then the following relations hold:

$$\boxed{V_t \leq V_0 e^{-\alpha t} + \frac{\beta}{\alpha}\left(1 - e^{-\alpha t}\right), \quad \limsup_{t\to\infty} V_t \leq \frac{\beta}{\alpha}}$$

Proof.

$$V_t = G_t + c, \; c = \frac{\beta}{\alpha}, \; \dot{G}_t \leq -\alpha\left(G_t + c\right) + \beta = -\alpha G_t,$$
$$G_t \leq G_0 e^{-\alpha t} \to V_t - c \leq \left(V_0 - c\right) e^{-\alpha t},$$
$$V_t \leq \left(V_0 - c\right) e^{-\alpha t} + c = V_0 e^{-\alpha t} + c\left(1 - e^{-\alpha t}\right).$$

Definition

- The set $\mathcal{E}(\bar{x}, P)$ of points x from R^n is referred to as the ellipsoid with the center in the point $\bar{x}$ and with the corresponding ellipsoidal matrix $P = P^{\intercal} \geq 0$ if for any $x \in \mathcal{E}(\bar{x}, P)$ the following inequality holds

$$(x - \bar{x})^{\intercal} P (x - \bar{x}) \leq 1. \tag{1}$$

- If $\bar{x} = 0$, then the ellipsoid $\mathcal{E}(P)$ is called the **central ellipsoid**, any point of which satisfies

$$x^{\intercal} P x \leq 1.$$

The semi - axis $r_i(P)$ of the ellipsoid $\mathcal{E}(\bar{x}, P)$ (or $\mathcal{E}(P)$) are equal to

$$\boxed{r_i(P) = \frac{1}{\sqrt{\lambda_i(P)}} \; (i = 1, ..., n) \, .} \tag{2}$$

If all $r_i(P) < \infty$, or equivalently, all $\lambda_i(P) > 0$ $(i = 1, ..., n)$, then such ellipsoid is named *Bodily ellipsoid*.
Obviously that an ellipsoid $\mathcal{E}(\bar{x}, P_1)$ is upload inside of an ellipsoid $\mathcal{E}(\bar{x}, P_2)$, that is,

$$\mathcal{E}(\bar{x}, P_1) \subset \mathcal{E}(\bar{x}, P_2), \tag{3}$$

if its semi - axis $r_i(P_1)$ are less then the corresponding semi-axis $r_i(P_2)$ of another ellipsoid $r_i(P_1) < r_i(P_2)$ $(i = 1, ..., n)$, that equivalently can be expressed as $\lambda_i(P_1) > \lambda_i(P_2)$ $(i = 1, ..., n)$,or as

$$\begin{array}{rl} P_1 > P_2 & \quad (P_1 - P_2 > 0) \, , \\ P_1^{-1} < P_2^{-1} & \quad (P_1^{-1} - P_2^{-1} < 0) \, . \end{array} \tag{4}$$

Definition of an attractive ellipsoid

Definition

The ellipsoid

$$\boxed{\mathcal{E}_{\mathring{x}}(P_{attr}) := \{x \in \mathbb{R}^n : (x - \mathring{x})^\intercal P_{attr} (x - \mathring{x}) \leq 1\}} \qquad (5)$$

with the center in the point $\mathring{x}$ and the ellipsoidal matrix $P_{attr} = P_{attr}^\intercal > 0$ is said to be **attractive** for some dynamic system if for any trajectories $\{x(t)\}_{t \geq 0}$ of this system

$$\boxed{\limsup_{t \to \infty} \; (x(t) - \mathring{x})^\intercal P_{attr} (x(t) - \mathring{x}) \leq 1.} \qquad (6)$$

Notice that if the attractive ellipsoid $\mathcal{E}_{\mathring{x}}(P)$ is located in the origine than $\mathring{x} = 0$, then (5) becomes

$$\boxed{\limsup_{t \to \infty} x(t)^\intercal P_{attr} x(t) \leq 1} \qquad (7)$$

and

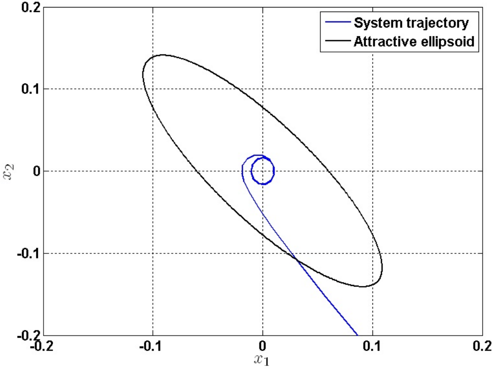

Figure 2: Two dimensional ellipsoid (ellipse).

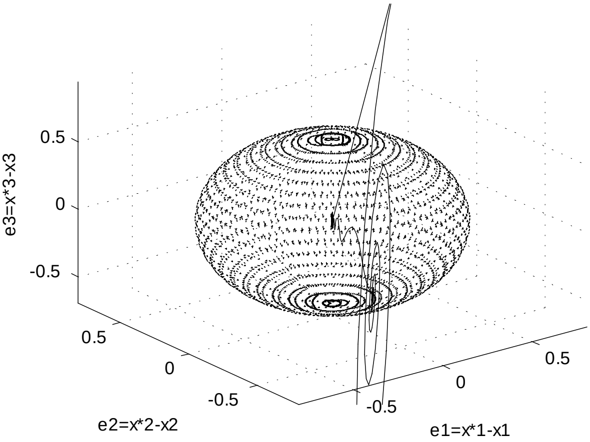

Figure 3: Three-dimensional ellipsoid.

Definition

Tracking error Δ_t is defined as

$$\boxed{\Delta_t := \hat{x}_t - x_t}$$

ODE for the tracking error is as follows

$$\dot{\Delta}_t = [A\hat{x}_t + Bu_t + L[y_t - C\hat{x}_t] + W_{0,t}\varphi(\hat{x}_t) + W_{1,t}\psi(\hat{x}_t) u_t] - \left[Ax_t + Bu_t + \tilde{\xi}_t\right],$$

or, after simplification,

$$\boxed{\dot{\Delta}_t = (A - LC)\Delta_t + L\eta_t + W_{0,t}\varphi(\hat{x}_t) + W_{1,t}\psi(\hat{x}_t) u_t - \tilde{\xi}_t} \quad (8)$$

Storage (or energetic) function

Definitions

The function $V_t = V\left(\Delta_t, W_{0,t}, W_{1,t}\right)$ equal

$$\boxed{\begin{array}{c} V_t = \Delta_t^{\intercal} P \Delta_t + \dfrac{k_0}{2} \mathrm{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \left(W_{0,t} - W_0^*\right) \\ \\ + \dfrac{k_1}{2} \mathrm{tr}\left(W_{0,t} - W_1^*\right)^{\intercal} \Lambda_1 \left(W_{0,t} - W_1^*\right), \end{array}} \tag{9}$$

with

$$\boxed{\begin{array}{c} P = P^{\intercal} > 0,\ \ \Lambda_0 = \Lambda_0^{\intercal} > 0,\ \ \Lambda_1 = \Lambda_1^{\intercal} > 0, \\ k_0 > 0,\ \ k_1 > 0,\ \ W_0^* \in R^{n \times k_{\varphi}},\ \ W_1^* \in R^{n \times k_{\psi}} \end{array}}$$

is referred to as the **Storage** (or **Energetic**) function.

Notice that V_t is the *Lyapunov-like function*, but not an exact Lyapunov function

Lie derivative of Storage function

Calculating the time derivative of V_t on the trajectories of ODE (8) (or Lie derivative) we get

$$\begin{array}{c} \dot{V}_t = 2\Delta_t^{\intercal} P \dot{\Delta}_t + k_0 \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \dot{W}_{0,t} + k_1 \operatorname{tr}\left(W_{1,t} - W_1^*\right)^{\intercal} \Lambda_1 \dot{W}_{1,t} \\ = 2\Delta_t^{\intercal} P \left[(A - LC)\Delta_t + L\eta_t + W_{0,t}\varphi(\hat{x}_t) + W_{1,t}\psi(\hat{x}_t) u_t - \tilde{\xi}_t \right] \\ + k_0 \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \dot{W}_{0,t} + k_1 \operatorname{tr}\left(W_{1,t} - W_1^*\right)^{\intercal} \Lambda_1 \dot{W}_{1,t} \end{array}$$

or equivalently (adding and substracting $\pm\alpha V_t$)

$$\begin{array}{c} \dot{V}_t = -\alpha V_t + 2\Delta_t^{\intercal} P \left(\frac{\alpha}{2} I_{n\times n} + A - LC \right) \Delta_t + 2\Delta_t^{\intercal} P L \eta_t + \\ 2\Delta_t^{\intercal} P \left[\left(W_{0,t} - W_0^*\right) \varphi(\hat{x}_t) + \left(W_{0,t} - W_1^*\right) \psi(\hat{x}_t) u_t \right] + \\ 2\Delta_t^{\intercal} P \left[W_0^* \varphi(\hat{x}_t) + W_1^* \psi(\hat{x}_t) u_t \right] - 2\Delta_t^{\intercal} P \tilde{\xi}_t \\ \alpha \frac{k_0}{2} \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \left(W_{0,t} - W_0^*\right) + k_0 \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \dot{W}_{0,t} \\ + \frac{k_1}{2} \operatorname{tr}\left(W_{0,t} - W_1^*\right)^{\intercal} \Lambda_1 \left(W_{0,t} - W_1^*\right) + k_1 \operatorname{tr}\left(W_{1,t} - W_1^*\right)^{\intercal} \Lambda_1 \dot{W}_{1,t} \end{array}$$

Quadratic form representation

The last relation can be expressed as

$$\dot{V}_t = -\alpha V_t +$$

$$\begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t)\, u_t \\ \tilde{\xi}_t \end{pmatrix}^{\intercal} \underbrace{\begin{bmatrix} \begin{matrix} P(A_\alpha\text{-}LC) + \\ (A_\alpha\text{-}LC)^{\intercal} P \end{matrix} & PL & PW_0^* & PW_1^* & P \\ L^{\intercal} P & 0 & 0 & 0 & 0 \\ (W_0^*)^{\intercal} P & 0 & 0 & 0 & 0 \\ (W_1^*)^{\intercal} P & 0 & 0 & 0 & 0 \\ P & 0 & 0 & 0 & 0 \end{bmatrix}}_{S_0} \begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t)\, u_t \\ \tilde{\xi}_t \end{pmatrix}$$

$$+\, 2\Delta_t^{\intercal} P \left[(W_{0,t} - W_0^*)\, \varphi(\hat{x}_t) + (W_{0,t} - W_1^*)\, \psi(\hat{x}_t)\, u_t \right] +$$

$$\alpha \frac{k_0}{2} \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \left(W_{0,t} - W_0^*\right) + k_0 \operatorname{tr}\left(W_{0,t} - W_0^*\right)^{\intercal} \Lambda_0 \dot{W}_{0,t}$$

$$+\, \alpha \frac{k_1}{2} \operatorname{tr}\left(W_{1,t} - W_1^*\right)^{\intercal} \Lambda_1 \left(W_{1,t} - W_1^*\right) + k_1 \operatorname{tr}\left(W_{1,t} - W_1^*\right)^{\intercal} \Lambda_1 \dot{W}_{1,t}$$

where

$$A_\alpha := \frac{\alpha}{2} I_{n \times n} + A$$

Representation as a trace

Let us use the identity

$$
\boxed{
\begin{array}{c}
2\Delta_t^{\intercal} P \left[\left(W_{0,t} - W_0^* \right) \varphi \left(\hat{x}_t \right) + \left(W_{0,t} - W_1^* \right) \psi \left(\hat{x}_t \right) u_t \right] = \\
2 \left[\left(W_{0,t} - W_0^* \right) \varphi \left(\hat{x}_t \right) + \left(W_{0,t} - W_1^* \right) \psi \left(\hat{x}_t \right) u_t \right]^{\intercal} P \Delta_t = \\
2 \left[\varphi^{\intercal} \left(\hat{x}_t \right) \left(W_{0,t} - W_0^* \right)^{\intercal} + u_t^{\intercal} \psi^{\intercal} \left(\hat{x}_t \right) \left(W_{0,t} - W_1^* \right)^{\intercal} \right] P \Delta_t = \\
\operatorname{tr} \left\{ 2 \left[\varphi^{\intercal} \left(\hat{x}_t \right) \left(W_{0,t} - W_0^* \right)^{\intercal} + u_t^{\intercal} \psi^{\intercal} \left(\hat{x}_t \right) \left(W_{0,t} - W_1^* \right)^{\intercal} \right] P \Delta_t \right\} \\
= \operatorname{tr} \left\{ \left(W_{0,t} - W_0^* \right)^{\intercal} \left[2 P \Delta_t \varphi^{\intercal} \left(\hat{x}_t \right) \right] \right\} \\
+ \operatorname{tr} \left\{ \left(W_{1,t} - W_1^* \right)^{\intercal} \left[2 P \Delta_t u_t^{\intercal} \psi^{\intercal} \left(\hat{x}_t \right) \right] \right\}
\end{array}
}
$$

Combining the trace terms together we obtain

$$\dot{V}_t = -\alpha V_t + \begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t) u_t \\ \tilde{\xi}_t \end{pmatrix}^{\intercal} S_0 \begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t) u_t \\ \tilde{\xi}_t \end{pmatrix} + \text{Learn}_0 + \text{Learn}_1,$$

where

$$\text{Learn}_0 := \operatorname{tr}\left\{ (W_{0,t} - W_0^*)^{\intercal} \left[2P\Delta_t \varphi^{\intercal}(\hat{x}_t) + \alpha \frac{k_0}{2} \Lambda_0 (W_{0,t} - W_0^*) + k_0 \Lambda_0 \dot{W}_{0,t} \right] \right\},$$

$$\text{Learn}_1 := \operatorname{tr}\left\{ (W_{1,t} - W_1^*)^{\intercal} \left[2P\Delta_t u_t^{\intercal} \psi^{\intercal}(\hat{x}_t) + \alpha \frac{k_1}{2} \Lambda_1 (W_{1,t} - W_1^*) + k_1 \Lambda_1 \dot{W}_{1,t} \right] \right\}.$$

Negative quadratic form

Let us use the following representation

$$\begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t) u_t \\ \tilde{\xi}_t \end{pmatrix}^{\intercal} \underbrace{\begin{bmatrix} S_{0,11} & PL & PW_0^* & PW_1^* \\ L^{\intercal} P & 0 & 0 & 0 \\ (W_0^*)^{\intercal} P & 0 & 0 & 0 \\ (W_1^*)^{\intercal} P & 0 & 0 & 0 \end{bmatrix}}_{S_0} \begin{pmatrix} \Delta_t \\ \eta_t \\ \varphi(\hat{x}_t) \\ \psi(\hat{x}_t) u_t \\ \tilde{\xi}_t \end{pmatrix} =$$

$$z_t^{\intercal} \underbrace{\begin{bmatrix} S_{0,11} & PL & PW_0^* & PW_1^* \\ L^{\intercal} P & -\varepsilon I_{m \times m} & 0 & 0 \\ (W_0^*)^{\intercal} P & 0 & -\varepsilon I_{k_\varphi \times k_\varphi} & 0 \\ (W_1^*)^{\intercal} P & 0 & 0 & -\varepsilon I_{k_\psi \times k_\psi} \end{bmatrix}}_{S_\varepsilon} z_t$$

$$+\varepsilon \left(\|\eta_t\|^2 + \|\varphi(\hat{x}_t)\|^2 + \|\psi(\hat{x}_t) u_t\|^2 + \|\tilde{\xi}_t\|^2 \right),$$

where

$$S_{0,11} = P(A_\alpha - LC) + (A_\alpha - LC)^{\intercal} P, \; z_t := \left(\Delta_t^{\intercal}, \eta_t^{\intercal}, \varphi^{\intercal}(\hat{x}_t), [\psi(\hat{x}_t) u_t]^{\intercal}\right)^{\intercal}.$$

Negative quadratic form

Using the upper estimates

$$\left\|\varphi\left(x\right)\right\| \leq \varphi_{+}, \ \left\|\psi\left(x\right)\right\| = \lambda_{\max}^{1/2}\left(\psi\left(x\right)^{\intercal}\psi\left(x\right)\right) \leq \psi_{+}, \ \left\|u_t\right\| \leq k,$$

we get

$$\begin{array}{l}\left\|\eta_t\right\|^2 + \left\|\varphi\left(\hat{x}_t\right)\right\|^2 + \left\|\psi\left(\hat{x}_t\right)u_t\right\|^2 + \left\|\tilde{\xi}_t\right\|^2 \leq \\ \underbrace{\eta_+^2 + \varphi_+^2 + \psi_+^2 k^2 + c_0 + c_1\left(d_0 + d_1 k\right)^2}_{\beta} = \beta\end{array}$$

and finally

$$\dot{V}_t \leq -\alpha V_t + z_t^{\intercal} S_{\varepsilon} z_t^{\intercal} + \mathrm{Learn}_0 + \mathrm{Learn}_1 + \varepsilon\beta,$$

Main result on Attractive Ellipsoid

Now we are ready to formulate the main result.

Theorem

If under the accepted assumption (on the upper bounds) there exist marices $P > 0$, A, L, W_0^*, W_1^*, *and constants* $\alpha > 0$, $\varepsilon > 0$, *such the matrix* $S_{\alpha,\varepsilon}$ *is strictly negative, that is,*

$$S_{\alpha,\varepsilon} := \begin{bmatrix} \begin{array}{c} P\left(\frac{\alpha}{2}I_{n\times n}+A-LC\right)+ \\ \left(\frac{\alpha}{2}I_{n\times n}+A-LC\right)^{\intercal}P \end{array} & PL & PW_0^* & PW_1^* \\ L^{\intercal}P & -\varepsilon I_{m\times m} & 0 & 0 \\ \left(W_0^*\right)^{\intercal}P & 0 & -\varepsilon I_{k_\varphi\times k_\varphi} & 0 \\ \left(W_1^*\right)^{\intercal}P & 0 & 0 & -\varepsilon I_{k_\psi\times k_\psi} \end{bmatrix} < 0 \tag{10}$$

Main result on Attractive Ellipsoid (continuation)

Theorem (continuation)

and Learning Laws

$$\begin{aligned} \dot{W}_{0,t} &= -\frac{\alpha}{2}\left(W_{0,t} - W_0^*\right) - 2\frac{\Lambda_0^{-1}}{k_0} P \Delta_t \varphi^{\intercal}\left(\hat{x}_t\right), \\ \dot{W}_{1,t} &= \frac{\alpha}{2}\left(W_{1,t} - W_1^*\right) - 2\frac{\Lambda_1^{-1}}{k_1} P \Delta_t u_t^{\intercal} \psi^{\intercal}\left(\hat{x}_t\right), \end{aligned} \tag{11}$$

hold, then the storage function V_t satisfies the following ODE

$$\dot{V}_t \leq -\alpha V_t + \varepsilon\beta, \tag{12}$$

and

$$\limsup_{t\to\infty} V_t \leq \varepsilon\frac{\beta}{\alpha} \tag{13}$$

Ellipsoidal matrix

In view of the relation

$$\boxed{\Delta_t^\intercal P \Delta_t \leq V_t = \Delta_t^\intercal P \Delta_t + \sum_{i=0}^{1} \frac{k_i}{2} \operatorname{tr}\left(W_{i,t} - Wi\right)^\intercal \Lambda_i \left(W_{i,t} - Wi\right)}$$

we may conclude that

$$\limsup_{t \to \infty} \Delta_t^\intercal P \Delta_t \leq \varepsilon \frac{\beta}{\alpha},$$

or equivalently

$$\boxed{\limsup_{t \to \infty} \Delta_t^\intercal \left(\frac{\alpha}{\varepsilon \beta} P \right) \Delta_t \leq 1.}$$

Fact

Attractive ellipsoid $\mathcal{E}_0(P_{attr})$ is defined by the matrix

$$\boxed{P_{attr} = \frac{\alpha}{\varepsilon \beta} P}$$