Scaled Relative Graphs of Data-Driven Systems

Talitha Nauta

I. Introduction

Machine learning models, such as neural networks and other data-driven approaches, are being used in a wider area than ever before. This includes safety critical systems in uncertain and unstructured environments. As machine learning models are complex black-box models, few tools exist to perform safety and performance analysis for these systems. A method that can help bridge this gap is the Scaled Relative Graph (SRG). SRGs provide an interpretable tool to analyse complex systems. In this article, we explain what the SRG is and how we can draw SRGs for data-driven systems.

II. Scaled Relative Graphs

The SRG of a linear operator T, denoted SRG(T), is a subset of the complex plane and is defined as follows [2]

$$\left\{\frac{\|y\|}{\|x\|}\exp\left(\pm i\arccos\left(\frac{\text{Re}(\langle y,x\rangle)}{\|y\|\|x\|}\right)\right): x\in\mathcal{H}, y=Tx\right\}. \quad (1)$$

From the definition it follows that the SRG captures some geometric features of the input-output pairs of the operator T. The first term $\frac{\|y\|}{\|x\|}$ represents the scaling of the output relative to the input and the exponent captures the angle between the input and output, as the angle θ between $x \in \mathcal{H}$ and $y \in \mathcal{H}$ is usually defined as

$$\cos\theta = \frac{\text{Re}(\langle y,x\rangle)}{\|y\|\|x\|}.$$

Note that the SRG is mirrored in the real axis as it captures both positive and negative angles.

Some conclusions that can be drawn from the SRG are for example that the operator T is contractive if the SRG lies inside the unit circle, as $\|y\|/\|x\| \leq 1$ in that case. Also, if the SRG is a subset of the right half plane the operator T is passive, as it corresponds to $\text{Re}(\langle y,x\rangle) > 0$.

III. SRGs from state space representation

We consider a discrete time invariant system with input $u(k) \in \mathbb{R}$, output $y(k) \in \mathbb{R}$ and internal state $x(k) \in \mathbb{R}^n$

$$\mathcal{G}: \begin{cases} x(k+1) = A\,x(k) + B\,u(k) \\ y(k) = C\,x(k) + D\,u(k) \end{cases}. \quad (2)$$

This system represents a transfer function between in- and output given by $G(s) = C(sI-A)^{-1}B + D$. To draw the SRG($\mathcal{G}$) of we need to use some properties of the SRG. First, note that SRG($T - \alpha I$) = SRG(T) $- \alpha$. Next, denote the

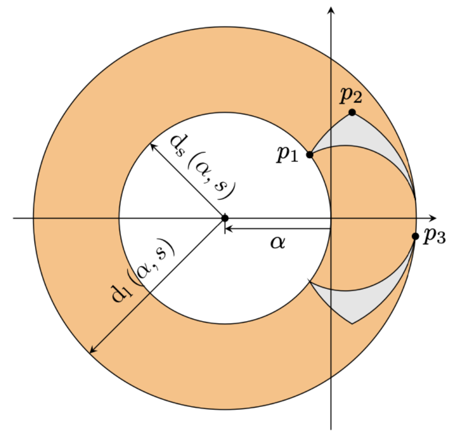

Fig. 1. The shortest and longest distance from the point α to the set defined by the grey area. The grey set must therefore lie inside the orange annulus.

shortest and the longest distance from a point $\gamma \in \mathbb{C}$ to a set $s \subseteq \mathbb{C}$ as

$$d_s(\gamma, s) = \inf\{|z-\gamma| : z \in s\}$$
$$d_l(\gamma, s) = \sup\{|z-\gamma| : z \in s\}.$$

Then it follows from the definition of the SRG that

$$d_s(\alpha, \text{SRG}(\mathcal{G})) = \sigma_{\text{min}}(G - \alpha I)$$
$$d_l(\alpha, \text{SRG}(\mathcal{G})) = \sigma_{\text{max}}(G - \alpha I),$$

where σ denotes the singular value. An example of this can be see in Figure 1, where the grey area corresponds to the SRG. The maximum singular value is given by $\sigma_{\text{max}}(G) = \|G\|_{L_\infty}$. The L_∞ norm of the system can be found solving the following LMI [3]

$$\begin{aligned} &\min \gamma \\ &\text{ s.t. } P = P^T \\ &\begin{bmatrix} A^TPA - P + C^TC & A^TPB + C^TD \\ B^TPA + D^TC & B^TPB + D^TD - \gamma^2 I \end{bmatrix} \preceq 0 \end{aligned} \quad (3)$$

where $\arg\min\gamma = \sigma_{\text{max}}(G)$. The minimum singular value is given by $\sigma_{\text{min}} = 1/\|G^{-1}\|_{L_\infty}$. If the inverse of the system does not exist $\sigma_{\text{min}} = 0$. If we solve this LMI for a grid of $\alpha \in \mathbb{R}$ we get an annulus containing the SRG for every α. If the range of α is wide enough the intersection of these annuli gives an outer approximation of the SRG.

We illustrate the results with an example. Take the following system

$$P: \begin{cases} x(k+1) = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.9 \end{bmatrix} x(k) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(k) \\ y(k) = \begin{bmatrix} 0.2 & 0.3 \end{bmatrix} x(k) + 1\,u(k) \end{cases}. \quad (4)$$

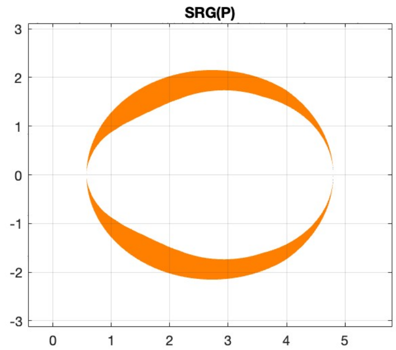

Fig. 2. The SRG of the system P in (4).

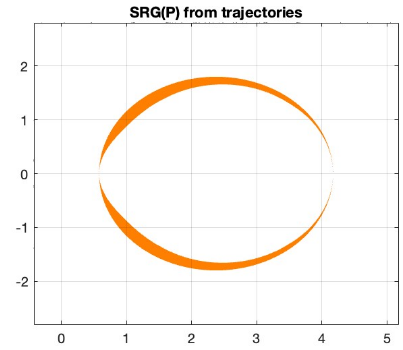

Fig. 3. The SRG of the system P in (4) reconstructed by input output trajectories.

Then the resulting SRG can be seen in Figure 2.

IV. From data trajectories to state space

For a linear time invariant data driven system the matrices A, B, C, D that define (2) are unknown. They can however be reconstructed from persistently exciting input-output trajectories and the SRG can be drawn using the reconstructed state space model. As the state space model of a system is not unique, the reconstructed system $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ might not be equivalent, but it will be zero-state equivalent and represent the same transfer function.

We define $U_{i,j} = \begin{bmatrix} u(i) & u(i+1) \dots u(j-1) \end{bmatrix}$ and $Y_{i,j} = \begin{bmatrix} y(i) & y(i+1) \dots y(j-1) \end{bmatrix}$ for $j > i$. Given an input trajectory $U_{-2,T}$ and an output trajectory $Y_{-2,T}$ we construct two matrices

$$\hat{X} = \begin{bmatrix} U_{0,T} \\ Y_{-2,T-2} \\ U_{-2,T-2} \end{bmatrix} \text{ and } \hat{Y} = \begin{bmatrix} Y_{-1,T-1} \\ U_{-1,T-1} \end{bmatrix} \tag{5}$$

If rank$(\hat{X}) = 2n+1$ the system is persistently exciting, which requires $T \geq 2n + 1$, the state space representation can be reconstructed as follows [1]

$$\begin{bmatrix} \tilde{B} & \tilde{A} \end{bmatrix} = \hat{X}\hat{Y}^{\dagger} \tag{6}$$

$$\tilde{C} = e_n^{\top} \hat{X}\hat{Y}^{\dagger} \begin{bmatrix} 0_{1\times 2n} \\ I_{2n} \end{bmatrix} \tag{7}$$

$$\tilde{D} = e_n^{\top} \hat{X}\hat{Y}^{\dagger} e_1. \tag{8}$$

Where $\dagger$ denotes the pseudo inverse. Note that this realisation is of order $2n$ and therefore non-minimal.

From this state-space representation we can draw the SRG by solving the LMI in (3). An example of this can be seen in Figure 3. Here we reconstruct the SRG from (4) using input-output trajectories of length $T = 50$ where the input is a random normal distributed signal, with zero mean and unit variance, and additional normal distributed output noise with variance 10^{-3}. This SRG is very similar to the SRG of the original system. It is lies in the same area of the complex plane but it is slightly more narrow.

References

[1] Claudio De Persis and Pietro Tesi. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3):909–924, 2020.

[2] Richard Pates. The scaled relative graph of a linear operator, 2021.

[3] Anders Rantzer. On the kalman—yakubovich—popov lemma. *Systems Control Letters*, 28(1):7–10, 1996.