En introduktion till lärande reglering

Övningar

Martin Enqvist

2020-05-18

Uppgift RL1

Betrakta ett system med en skalär insignal och ett skalärt tillstånd och antag att man använder förstärkningsinlärning och har skattat $\mathcal{Q}$-funktionen

$$\mathcal{Q}(x_k, u_k) = 0.5(1.2x_k^2 + 0.9x_k u_k + 0.6u_k^2)$$

för en viss tillståndsåterkoppling. Vad blir den uppdaterade tillståndsåterkopplingen $L_{ny} = S_{uu}^{-1} S_{ux}$?

Uppgift RL2

(a)

Matrisen P_L som ingår i värdefunktionen $V_L(x_k)$ är definierad som

$$P_L = \sum_{i=0}^{\infty} (A - BL)^{i,T}(Q + L^T RL)(A - BL)^i$$

för en viss tillståndsåterkoppling L. Ta fram ett uttryck för

$$(A - BL)^T P_L (A - BL) - P_L$$

genom att sätta in definitionen av P_L.

(b)

Betrakta ett system

$$\begin{aligned} \dot{x} &= ax + bu \\ y &= x \end{aligned}$$

med en skalär insignal och ett skalärt tillstånd. Använd resultatet från a-uppgiften för att ta fram ett uttryck för P_L i detta specialfall.

(c)

Antag att $a = 0.9$, $b = 1$, $Q = 2$ och $R = 1$ och beräkna den optimala tillståndsåterkopplingen

$$L_{LQ} = (B^T PB + R)^{-1} B^T PA$$

där P är den positiva semidefinita lösningen till den tidsdiskreta algebraiska riccatiekvationen

$$A^T PA - P + Q - A^T PB(B^T PB + R)^{-1} B^T PA = 0.$$

(d)

Betrakta samma system som i c-uppgiften och antag att den initiala tillståndsåterkopplingen är $L^0 = 0.2$. Om modellen är känd kan den uppdaterade tillståndsåterkopplingen beräknas som $L^{j+1} = (B^T P_{L^j} B + R)^{-1} B^T P_{L^j} A$. Använd detta uttryck för att beräkna de tre första uppdateringarna L^1, L^2 och L^3 av tillståndsåterkopplingen som man skulle få om man använde förstärkningsinlärning för det här systemet.

Förberedelse inför uppgift RL3

Matlabfunktionen `rllqr` implementerar förstärkningsinlärning för linjärkvadratisk reglering (LQ) av ett andra ordningens system med en insignal. Funktionen anropas med

```
rllqr(Acont,Bcont,Ts,x1,Q,R,L0,sigma_e,Nls,N)
```

där `Acont` och `Bcont` definierar det kontinuerliga systemets tillståndsbeskrivning, `Ts` är sampeltiden och `x1` är initialtillståndet x_1. Vidare är `Q` och `R` matriserna som definierar målfunktionen och `L0` den initiala tillståndsåterkopplingen. Parametern `sigma_e` definierar den normalfördelade störsignalen e_k:s standardavvikelse, `Nls` är storleken på minstakvadratproblemen som löses i algoritmen och `N` är det totala antalet sampel i experimentet. Se även hjälptexten till funktionen via `help rllqr`.

Uppgift RL3

(a)

Simulera experiment med $N = 200$ sampel där förstärkningsinlärning används på tre olika linjära system

$$\dot{x}(t) = \begin{pmatrix} -2 & -2 \\ 2 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u(t), \tag{1}$$

$$\dot{x}(t) = \begin{pmatrix} -0.5 & -1 \\ 0.4 & -1 \end{pmatrix} x(t) + \begin{pmatrix} 1 \\ -0.1 \end{pmatrix} u(t), \tag{2}$$

och

$$\dot{x}(t) = \begin{pmatrix} -5 & 80 \\ 1 & -10 \end{pmatrix} x(t) + \begin{pmatrix} -1 \\ 1 \end{pmatrix} u(t). \tag{3}$$

Antag att $T_S = 0.5\,\mathrm{s}$, $x_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ och att vi väljer $Q = I$ och $R = 1$ samt $L^0 = \begin{pmatrix} 1 & 1 \end{pmatrix}$, $\sigma_e = 0.01$ och $N_{LS} = 20$ för alla systemen. Konvergerar regulatorparametrarna i de olika fallen och till vilka värden?

(b)

Undersök nu hur det blir med samma sampeltid, initialtillstånd, Q, R samt parametrar i algoritmen för förstärkningsinlärning om systemet istället är

$$\dot{x}(t) = \begin{pmatrix} -0.2 & -2 \\ 20 & 0 \end{pmatrix} x(t) + \begin{pmatrix} -10 \\ 0 \end{pmatrix} u(t). \tag{4}$$

Kan du hitta vad som är problemet? Testa även fallet när den initiala tillståndsåterkopplingen är noll. Varför skulle det kunna vara ett lämpligt val?

Svar till RL1

Identifiering av termer i $\mathcal{Q}$-funktionen ger $S_{uu} = 0.6$ och $S_{ux} = 0.45$, vilket ger $L_{ny} = 0.75$.

Svar till RL2

(a)

$$(A - BL)^T P_L (A - BL) - P_L = -Q - L^T RL$$

(b)

$$P_L = \frac{-Q - L^T RL}{(a - bL)^2 - 1}$$

(c)

Insättning av de numeriska värdena i riccatiekvationen och multiplikation med $P + 1$ ger ekvationen

$$(0.9^2 P - P + 2)(P + 1) - 0.9^2 P^2 = -P^2 + 1.81P + 2 = 0,$$

som har (den positiva) lösningen $P = 2.584$. Detta ger $L_{LQ} = \frac{0.9P}{P+1} \approx 0.6489$.

(d)

Iterationer med uttrycket från b-uppgiften och uttrycket för L^{j+1} i uppgiften ger $L^1 = 0.7200$, $L^2 = 0.6502$ och $L^3 = 0.6489$. Resultatet L^3 av den tredje iteration är mycket nära det optimala L_{LQ}.

Svar till RL3

(a)

Regulatorparametrarna l_1 och l_2 i $L = \begin{pmatrix} l_1 & l_2 \end{pmatrix}$ konvergerar för alla tre systemen trots att vi inte utnyttjar någon modellinformation och använder en identisk version av algoritmen för förstärkningsinlärning i de tre fallen. Parametern l_1 konvergerar till 0.26, 0.41 respektive 0.51 och l_2 konvergerar till 0.031, -0.24 respektive 3.48.

(b)

Förstärkningsinlärningsalgoritmen konvergerar inte för det här systemet och tillstånden uppvisar ett instabilt beteende. Problemet är att den initiala tillståndsåterkopplingen L^0 inte stabiliserar systemet. Eftersom det öppna systemet är stabilt fungerar det att sätta den initiala tillståndsåterkopplingen till noll. Med det valet konvergerar regulatorparametrarna men tillstånden får en oscillativ insvängning.