Βέλτιστος Έλεγχος Συστημάτων

Ενότητα 10: Δυναμικός προγραμματισμός

Καθηγητής Αντώνιος Αλεξανδρίδης

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

Σημείωμα Αδειοδότησης

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης creative commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκεινται σε άλλου τύπου άδειες χρήσης, άδεια αναφέρεται ρητώς.

Δυναμικός προγραμματισμός

Δυναμικός προγραμματισμός

Ας θεωρήσουμε το σύστημα διακριτού χρόνου, εν γένει πολλών εισόδων – πολλών εξόδων, που περιγράφεται από την εξίσωση διαφορών:

$$x(j+1) = F\left[x(j), u(j)\right] \qquad\qquad x(0) = x_0$$

και το κριτήριο κόσους δίνεται από τη σχέση:

$$J = \sum_{j=0}^{N-1} L_j\left[x(j), u(j)\right] + \Phi\left[x(N)\right]$$

Δυναμικός προγραμματισμός

Σύμφωνα με την αρχή της βελτιστοποίησης θα πρέπει για οποιαδήποτε χρονική στιγμή j=i και οποιαδήποτε κατάσταση *x(i)* η πολιτική ελέγχου από εκεί και πέρα να είναι βέλτιστη. Επομένως, με βάση το κριτήριο κόστους *J* ορίζουμε για κάθε *x(i)* και *i* την πολιτική βελτιστοποίησης με τη μαθηματική έκφραση:

$$V_i\left[x(i)\right] = \min_{u\langle i\rangle}\left\{\sum_{j=1}^{N-1} L_j\left[x(j),u(j)\right] + \Phi\left[x(N)\right]\right\} \quad (10.1)$$

Δυναμικός προγραμματισμός

Σύμφωνα όμως με την Αρχή της Βελτιστοποίησης, έχουμε:

$$V_i\left[x(i)\right] = \min_{u\langle i\rangle}\left\{L_i\left[x(i),u(i)\right] + \sum_{j=i+1}^{N-1} L_j\left[x(j),u(j)\right] + \Phi\left[x(N)\right]\right\} \quad (10.2)$$

επειδή όμως προφανώς ισχύει:

$$V_{i+1}\left[x(i+1)\right] = \sum_{j=i+1}^{N-1} L_j\left[x(j),u(j)\right] + \Phi\left[x(N)\right] \quad (10.3)$$

Δυναμικός προγραμματισμός

η σχέση (10.1) καταλήγει στην :

$$\boxed{V_i\left[x(i)\right]=\min_{u\langle i\rangle}\left\{L_i\left[x(i),u(i)\right]+V_{i+1}\left[x(i+1)\right]\right\}} \qquad (10.4)$$

με τελική τιμή στο όριο i=N:

$$\boxed{V_N\left[x(N)\right]=\Phi\left[x(N)\right]} \qquad (10.5)$$

Δυναμικός προγραμματισμός

Η εξίσωση (10.4) ονομάζεται εξίσωση Hamilton-Jacobi-Bellman και αποτελεί έναν επαναληπτικό αλγόριθμο που ξεκινάει από το όριο i=N , με τιμή αυτή που ορίζεται από τη σχέση (10.5), και δίνει το βέλτιστο νόμο ελέγχου για συστήματα διακριτού χρόνου, σε κάθε βήμα, συναρτήσει του διανύσματος κατάστασης στο σημείο ελαχίστου της, δηλαδή εκεί όπου:

$$\frac{\partial}{\partial u(i)}\left\{L_i\left[x(i),u(i)\right]+V_{i+1}\left[x(i+1)\right]\right\}=0 \qquad (10.6)$$

Δυναμικός προγραμματισμός

Η διαδικασία αυτή εύρεσης του βέλτιστου ελέγχου σε κάθε βήμα για συστήματα διακριτού χρόνου με χρήση της εξίσωσης Hamilton-Jacobi-Bellman είναι ευρέως διαδεδομένη με την ονομασία <u>Δυναμικός Προγραμματισμός</u>.

$$\frac{\partial}{\partial u(i)}\left\{L_i\left[x(i),u(i)\right]+V_{i+1}\left[x(i+1)\right]\right\}=0 \qquad (10.6)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Ας θεωρήσουμε το γραμμικό σύστημα διακριτού χρόνου:

$$x(j+1) = Fx(j) + Gu(j) \tag{10.7}$$

με κριτήριο κόστους :

$$J = \sum_{j=0}^{N-1} \left[x^T(j+1)Qx(j+1) + \lambda u^2(j) \right] \tag{10.8}$$

Η εξίσωση Hamilton-Jacobi-Bellman θα είναι:

$$\begin{aligned} V_i\left[x(i)\right] &= \min_{u\langle i\rangle}\left\{V_{i+1}\left[x(i+1)\right] + x^T(i+1)Qx(i+1) + \lambda u^2(i)\right\} = \\ &= \min_{u\langle i\rangle}\left\{V_{i+1}\left[x(i+1)\right] + \left[Fx(i)+Gu(i)\right]^T Q\left[Fx(i)+Gu(i)\right] + \lambda u^2(i)\right\} \end{aligned} \tag{10.9}$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

- Ξεκινώντας από το τελευταίο βήμα , για <u>i=N-1</u> έχουμε την ακόλουθη έκφραση για την εξίσωση H-J-B :

$$V_{N-1}\left[x\left(N-1\right)\right]= \tag{10.10}$$

$$=\min_{u\langle N-1\rangle}\left\{V_N\left[x\left(N\right)\right]+\left[Fx\left(N-1\right)+Gu\left(N-1\right)\right]^T Q\left[Fx\left(N-1\right)+Gu\left(N-1\right)\right]+\lambda u^2\left(N-1\right)\right\}$$

όπου

$$V_N\left[x\left(N\right)\right]=0 \tag{10.11}$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Ο δε νόμος ελέγχου όπως προκύπτει από την εφαρμογή της σχέσης (10.6) για i=N-1 είναι:

$$\boxed{u(N-1) = -\frac{G^T QF}{G^T QG + \lambda} x(N-1)} \qquad (10.12)$$

αν ονομάσουμε το συντελεστή ανάδρασης στο βήμα i=N-1 σαν H(N-1) παίρνουμε:

$$H(N-1) = -\frac{G^T QF}{G^T QG + \lambda} \qquad (10.13)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Αν τώρα αντικαταστήσουμε το *u* στην αναδρομική σχέση (10.10) παίρνουμε:

$$V_{N-1}\left[x(N-1)\right] = x^T(N-1)P(N-1)x(N-1) \qquad (10.14)$$

όπου :

$$\begin{aligned} P(N-1) = {} & \left[F + GH(N-1)\right]^T Q\left[F + GH(N-1)\right] + \\ & + \lambda H^T(N-1)H(N-1) \end{aligned} \qquad (10.15)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

- Πηγαίνοντας στο προηγούμενο βήμα, για i=N-2, αντίστοιχα έχουμε:

$$V_{N-2}\left[x\left(N-2\right)\right]= \tag{10.16}$$

$$=\min_{u\langle N-2\rangle}\left\{V_{N-1}\left[x\left(N-1\right)\right]+\left[Fx\left(N-2\right)+Gu\left(N-2\right)\right]^{T}Q\left[Fx\left(N-2\right)+Gu\left(N-2\right)\right]+\lambda u^{2}\left(N-2\right)\right\}$$

όπου το $V_{N-1}[x(N-1)]$ είναι αυτό που ορίστηκε στο βήμα i=N-1 από την (10.14). Σε αυτή την περίπτωση ο νόμος ελέγχου προκύπτει από την (10.16), για i=N-2 , ότι είναι:

$$\boxed{u\left(N-2\right)=-\frac{G^{T}\left[Q+P\left(N-1\right)\right]F}{G^{T}\left[Q+P\left(N-1\right)\right]G+\lambda}x\left(N-2\right)} \tag{10.17}$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Ο συντελεστής ανάδρασης είναι:

$$H(N-2) = -\frac{G^T\left[Q+P(N-1)\right]F}{G^T\left[Q+P(N-1)\right]G+\lambda} \qquad (10.18)$$

και η τιμή στην αναδρομική συνάρτηση κόστους προκύπτει:

$$V(N-2) = x^T(N-2)P(N-2)x(N-2) \qquad (10.19)$$

με :

$$P(N-2) = \left[F+GH(N-2)\right]^T\left[P(N-1)+Q\right]\left[F+GH(N-2)\right]+\lambda H^T(N-2)H(N-2) \qquad (10.20)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

- Αν θεωρήσουμε τώρα το γενικό βήμα i=j (0≤j≤N-1) και υποθέσουμε σύμφωνα με τα προηγούμενα ότι:

$$V_{j+1}\left[x(j+1)\right] = x^T(j+1)P(j+1)x(j+1) \tag{10.21}$$

τότε η $V_j[x(j)]$ θα είναι σύμφωνα με τη σχέση (10.6) δίνει το νόμο ελέγχου στο βήμα j:

$$V_j\left[x(j)\right] = \tag{10.22}$$

$$= \min_{u\langle j\rangle}\left\{x^T(j+1)Px(j+1) + \left[Fx(j)+Gu(j)\right]^T Q\left[Fx(j)+Gu(j)\right] + \lambda u^2(j)\right\}$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Η ελαχιστοποίηση της (10.22) σύμφωνα με τη σχέση (10.6) δίνει το νόμο ελέγχου στο βήμα j:

$$\boxed{u(j)=H(j)x(j)} \qquad (10.23)$$

όπου :

$$\boxed{H(j)=-\frac{G^T\left[Q+P(j+1)\right]F}{G^T\left[Q+P(j+1)\right]G+\lambda}} \qquad (10.24)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Αντικαθιστώντας τις (10.23) και (10.24) στην (10.22) προκύπτει για την V_j :

$$V_j\left[x(j)\right] = x^T(j)P(j)x(j) \qquad (10.25)$$

με

$$\boxed{P(j) = \left[F + GH(j)\right]^T \left[P(j+1) + Q\right]\left[F + GH(j)\right] + \lambda H^T(j)H(j)} \qquad (10.26)$$

με αρχική τιμή στο τελικό όριο N:

$$\boxed{P(N) = 0} \qquad (10.27)$$

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

Με τη μέθοδο του δυναμικού προγραμματισμού στο πρόβλημα της γραμμικής τετραγωνικής ρύθμισης είδαμε ότι καταλήξαμε στις αναδρομικές σχέσεις (10.23), (10.24) και (10.26) που ισχύουν για κάθε j: $0 \le j \le N-1$ με αρχική συνθήκη στο τελικό όριο $j=N$ αυτή που δίνεται από τη σχέση (10.27).

Το γραμμικό τετραγωνικό πρόβλημα ρύθμισης για συστήματα διαριτού χρόνου με τη μέθοδο του Δυναμικού Προγραμματισμού

<u>Παρατηρήσεις</u>

1. Στα συστήματα συνεχούς χρόνου η εφαρμογή της θεωρίας αυτής μάλλον <u>περιπλέκει</u> τα πράγματα στη λύση μερικών διαφορικών εξισώσεων σε σχέση με την αρχικά αναλυθείσα, μέθοδο της Αρχής του Ελαχίστου με χρήση της Hamiltonian.

2. Στα συστήματα διακριτού χρόνου η εφαρμογή του δυναμικού προγραμματισμού <u>απλοποιεί</u> τα πράγματα αφού καταλήγει σε απλούς σχετικά αναδρομικούς τύπους, που δίνουν απευθείας το βέλτιστο νόμο ελέγχου.

3. Και στις δύο προαναφερθείσες περιπτώσεις η λύση που προκύπτει από εφαρμογή της θεωρίας Hamilton-Jacobi-Bellman καταλήγει σε ελέγχους <u>πάντοτε κλειστού βρόχου.</u>

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Τέλος Ενότητας

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Αλεξανδρίδης Αντώνιος 2015. Αλεξανδρίδης Αντώνιος. «Βέλτιστος Έλεγχος Συστημάτων . Δυναμικός προγραμματισμός». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: https://eclass.upatras.gr/courses/EE887/

Σημείωμα Χρήσης Έργων Τρίτων

Όλα τα σχήματα, οι εικόνες και τα γραφήματα που παρουσιάστηκαν σε αυτήν την ενότητα είναι από το βιβλίο << Βέλτιστος Έλεγχος Συστημάτων >>, Αντώνης Θ. Αλεξανδρίδης, εκδόσεις Πανεπιστημίου Πατρών.