Projeto de um Controlador Ótimo para Posicionamento de um Painel Fotovoltaico via Aprendizado por Reforço

Alam Pablo Silva Belfort(e-mail: alam.pablo@ufma.discente.br)
Yan Ferreira da Silva(e-mail: yan.ferreira@ufma.discente.br)
João Viana da Fonseca Neto(e-mail: joao.fonsenca@ufma.br)

Programa de Pós Graduação em Engenharia Elétrica, Universidade Federal do Maranhão, Av. dos Portugueses 1966, São Luís, MA, Brasil.

Abstract: With the increase in energy demands in the global electricity sector, the need to use renewable energies such as photovoltaics is clear. The generation of photovoltaic energy depends, among other factors, on the type of photovoltaic panel, the temperature in the module and solar irradiance, therefore to optimize energy generation it is essential to track the position of the module to favor the best generation angle depending on the point maximum power. In this work, an approach for applying reinforcement learning, in the form of Action Dependent Dynamic Heuristic Programming (ADHDP), is proposed for tuning a Discrete controller of the Linear Quadratic Tracker (DLQT) of a motor DC for positioning the panel to obtain the point maximum power, with the lowest possible control effort (energy consumption).

Resumo: Com o aumento das demandas de energia no setor elétrico mundial, é notório a necessidade do uso de energias renováveis, como a fotovoltaica. A geração de energia fotovoltaica depende, dentre outros fatores, do tipo de painel fotovoltaico, da temperatura no módulo e irradiância solar, portanto para otimizar a geração de energia é fundamental rastrear a posição do módulo para favorecer o melhor ângulo de geração em função do ponto de máxima potência. Neste trabalho é proposto uma abordagem de aplicação de aprendizado por reforço, na forma de Programação Dinâmica Heurística Dependente da Ação (ADHDP), para sintonia de um controlador Discreto do Rastreador Linear Quadrático (DLQT) de um motor CC para posicionamento do painel para obtenção do ponto de máxima potência, otimizando o esforço de controle possível (consumo de energia).

Keywords: ADHDP; Optimal Control; Motor DC; DLQT, Reinforcement Learning.

Palavras-chaves: ADHDP; Controle Ótimo; Motor CC; DLQT, Aprendizado por Reforço.

1. INTRODUÇÃO

O consumo de energia elétrica cresceu muito nos últimos anos devido à mudança no estilo de vida da população, desenvolvimento da tecnologia, sofisticação da indústria e aquecimento global. A energia elétrica proveniente de fontes de energias renováveis, como a solar, tem desenvolvido um papel crucial devido à sua geração ser limpa, no entanto, os módulos fotovoltaicos têm eficiência inferior a 20%, portanto é fundamental que a superfície do módulo e a direção dos raios solares sejam ortonormais, e para manter esta condição, é necessário movimentar o módulo rastreando a trajetória do sol (Rahul and Hariharan, 2024).
Nesse sentido pesquisas têm sido desenvolvidas nos últimos anos, um controlador integral baseado em LQR para posicionamento angular e simulação baseada em modelo para analisar a factibilidade de implementação na prática é apresentado por Mazumdar et al. (2015). No trabalho desenvolvido por Akbar Gumilang et al. (2021) o rastreamento da posição solar é realizado em duas direções (vertical e horizontal) utilizando LQR e observador de Luenberger. Em Pataro et al. (2023) é proposto um LQR com ação *Feedforward* (FF) com objetivo de melhorar o desempenho dos controladores ótimos. Um estudo para validar o rastreamento de painéis fotovoltaicos em um satélite é realizado em Bai et al. (2023). Uma abordagem baseada em ADHDP para otimização de uma casa inteligente com painéis fotovoltaicos, turbinas eólicas e uma bateria é introduzida por Sterling and Tyler (2018). Em de Moura et al. (2019) é realizada a união da metodologia de controle ótimo e ADHDP para sintonia do sistema de controle ótimo discreto (DLQR).
Vale ressaltar que a maioria dos trabalhos para controle da posição do painel utiliza o LQR como método de controle ótimo, que tem como principal desvantagem a dependência de um modelo linear preciso do sistema. Este trabalho tem como principal contribuição o uso de uma metodologia crítica adaptativa (ADHDP) para sintonia dos ganhos aplicados ao DLQT com ou sem ação integral, suprindo as carências supracitadas do LQR.
O sistema é composto por módulos eletrônicos do sistema embarcado, planta e periféricos como sensores, bateria e IHM (Interface Homem Máquina), que é uma interface gráfica de ajuste do método de controle, que se comunica

com o sistema embarcado. A Figura 1 exibe a planta do sistema, o atuador (motor CC), o sensor de posição e velocidade e o *smart wireless sensor* proposto por da Silva et al. (2021), que indica a posição de referência para o controlador.

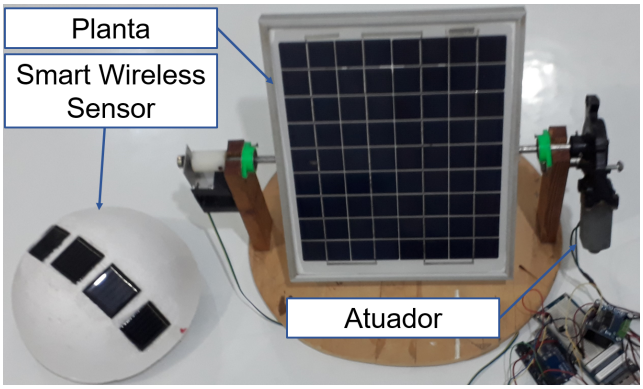

Figura 1. Planta do sistema composto por atuador, sensores, painel fotovoltaico e placas do sistema embarcado.

O trabalho está organizado em secções. Na Seção II são expostos os conceitos da modelagem do atuador. Na Seção III é exibida uma abordagem sobre as estratégias de controle utilizadas no projeto do DLQT e programação dinâmica aproximada, abordando conceitos sobre o princípio da otimalidade de *Bellman*, bem como a apresentação do algoritmo utilizado. Os procedimentos utilizados no experimento computacional em conjunto com os resultados são apresentados na Seção IV e a conclusão é apresentada na Seção V.

2. MODELO DO ATUADOR

Nesta Seção o modelo do atuador é apresentado conforme Gupta and Biswas (2023). O atuador que promove o movimento rotativo do painel fotovoltaico é o motor de corrente contínua (CC) de ímã permanente. Um conjunto de equações diferenciais caracterizam a dinâmica de funcionamento deste atuador, relacionando variáveis como velocidade, posição e corrente de armadura com a tensão aplicada.

Uma forma de representar a dinâmica do motor CC é a descrição em espaços de estados, dada por

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \tag{1}$$

e

$$y(t) = Cx(t), \tag{2}$$

sendo n a ordem do sistema, $x(t)$ o vetor de estados da planta de ordem $n \times 1$, A uma matriz $n \times n$, B a matriz de entrada $n \times 1$, $u(t)$ o sinal de controle e C a matriz de saída $1 \times n$. Essa descrição possibilita a aplicação de técnicas otimizadas para controle do motor, atendendo às métricas de desempenho estabelecidas.

No modelo representado em espaços de estados é possível verificar as características não-lineares variantes no tempo do motor CC como atrito, resposta da armadura e efeitos da saturação, segundo Gupta and Biswas (2023).

Para formulação em espaços de estados é definida a tensão de armadura V_a como entrada de controle $u(t)$. A relação dos estados é dada por

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} i_a(t) \\ \omega(t) \\ \theta(t) \end{bmatrix}, \tag{3}$$

sendo $i_a(t)$ a corrente de armadura, $\omega(t)$ a velocidade angular e $\theta(t)$ a posição do rotor. Considerando o circuito equivalente do motor CC, proposto em Sen (2021) e Ogata (2011), aplicando a lei de *Kirchhoff* das tensões e a lei de *Newton*. Tem-se as equações diferenciais ordinárias

$$\frac{di_a(t)}{dt} = -\frac{R_a}{L_a} i_a(t) - \frac{k\phi}{L_a}\omega(t) + \frac{V_a(t)}{L_a}, \tag{4}$$

e

$$\frac{d\omega(t)}{d} = \frac{k\phi}{J} i_a(t) - \frac{B}{J}\omega(t). \tag{5}$$

A Equação (4) rege o comportamento elétrico do motor CC, o comportamento mecânico é regido pela Equação (5) . Organizando na forma matricial

$$\begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \\ \dot{x_3}(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_a}{L_a} & -\frac{k\phi}{L_a} & 0 \\ \frac{k\phi}{J} & -\frac{B}{J} & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{L_a} \\ 0 \\ 0 \end{bmatrix} u(t), \tag{6}$$

$$y(t) = [0 \ 0 \ 1] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}. \tag{7}$$

sendo R_a e L_a a resistência e indutância da armadura, respectivamente, k é a constante da máquina, ϕ é o fluxo produzido pelo ímã permanente, J é a inércia do rotor e B representa o atrito do rotor. A Equação (7) mostra que a saída de interesse é o estado $x_3(t)$. O sistema representado na forma reduzida é dado por

$$\begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \end{bmatrix} = \begin{bmatrix} -\frac{R_a}{L_a} & -\frac{k\phi}{L_a} \\ \frac{k\phi}{J} & -\frac{B}{J} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{L_a} \\ 0 \end{bmatrix} u(t), \tag{8}$$

$$y(t) = [0 \ 1] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \tag{9}$$

Neste caso a saída de interesse é o estado $x_2(t)$.

3. PROJETO DLQT E PROGRAMAÇÃO DINÂMICA ADAPTATIVA

Nesta Seção são apresentados os projetos do seguidor linear quadrático sem ou com ação integral, conforme Ogata (2011) e de Moura et al. (2019). A programação dinâmica adaptativa, o princípio da otimalidade de *Bellman*, o método de sintonia do controlador e o algoritmo da ADHDP são apresentados conforme Di Ruscio (2012), de Moura et al. (2019), Lewis and Vrabie (2009) e Lewis et al. (2012).

3.1 Projeto do Seguidor Linear Quadrático sem ação integral - LQT

Para plantas do tipo 1 (com integrador), quando é necessário rastrear uma referência, o projeto de seguidor linear quadrático é suficiente para eliminar o erro de regime permanente. Considerando a função de transferência de um sistema do tipo 1, dada por

$$\frac{\theta(s)}{V_a(s)} = \frac{1}{s} \frac{k\phi}{(sL_a + R_a)(Js + B) + (k\phi)^2}. \tag{10}$$

Sendo $\theta(s)$ a saída. O sistema na representação em espaços de estados dada pelas Equações (6) e (7), com entrada de referência $r(t)$. O esforço de controle $u(t)$ é dado por

$$u(t) = -kx(t) + k_3 r(t), \tag{11}$$

sendo $x(t)$ o vetor de estados, k o vetor de ganhos da realimentação em espaços de estados, k_3 é o ganho aplicado ao erro $(r(t) - y(t))$ e $r(t)$ a transformada inversa de Laplace de uma posição de referência $\theta_{ref}(s)$ a ser seguida.

3.2 Projeto do Seguidor Linear Quadrático com ação Integral - LQT-I

Considerando a função de transferência de uma planta sem integrador (planta do tipo 0), dada por

$$\frac{\omega(s)}{V_a(s)} = \frac{k\phi}{(sL_a + R_a)(Js + B) + (k\phi)^2}. \tag{12}$$

A descrição em espaços de estados é exibida nas Equações (8) e (9), sendo $\omega(s)$ a saída. Para este caso é necessário inserir um integrador no ramo direto entre o comparador e a planta, como exibe a Figura 2. O esforço de controle é dado por

$$u(t) = -Kx(t) + k_I \xi, \tag{13}$$

sendo k_I o ganho integral ,K o ganho de realimentação dos estados e $r(t)$ a transformada inversa de Laplace de uma velocidade angular de referência $\omega_{ref}(s)$ a ser seguida. A relação entre a referência e saída é dada por

$$\dot{\xi}(t) = r(t) - y(t). \tag{14}$$

Considerando a Equação (14) com a nova variável do sistema ξ e a Equação (1). A nova descrição em espaço de estados é dada por

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\xi}(t) \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \xi(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} r(t). \tag{15}$$

O vetor de ganhos $\bar{K}$ dado por

$$\bar{K} = [\ K \mid -k_I\ ]. \tag{16}$$

Através do cálculo variacional encontra-se de forma *offline* os ganhos para as Equações (11) e (13) que garantem a otimalidade do sistema.

3.3 Princípio da Otimalidade de Bellman e Programação Dinâmica Adaptativa

Devido a aplicação em tempo real que utiliza microcontroladores no sistema embarcado de controle do processo é necessária uma análise do sistema em tempo discreto.
O princípio da otimalidade de *Bellman* constitui uma base sólida para programação dinâmica adaptativa e controle ótimo, dada por

$$V_g(x_j) = r(x_j, g(x_j)) + V_g(x_{j+1}), \tag{17}$$

sendo $V_g(x_j)$ a função de utilidade dada por

$$V_g(x_j) = F(x_j, g(x_j)), \tag{18}$$

sendo F o mapeamento que depende do estado e tem g como função de decisão. É necessário conhecer a política ótima no tempo $j+1$ para determinar a política ótima no tempo j. O princípio da otimalidade de *Bellman* consiste em um procedimento de retroação no tempo para resolver o problema de controle ótimo. Uma política ótima tem a propriedade que, independentemente das decisões anteriores, as próximas decisões devem ter política ótima em relação ao estado resultante dessas decisões anteriores (Lewis et al., 2012).
Em sistemas não lineares, os estados e a equação de custo são complexas para se resolver, tendo como agravante as restrições nas variáveis de controle. Neste contexto, segundo Lewis et al. (2012) e Lewis and Vrabie (2009), a programação dinâmica adaptativa é uma ferramenta imprescindível em razão de encontrar a solução de forma mais simples do que a abordagem do cálculo variacional.
Considerando o sistema não linear com representação dos estados discretos no tempo

$$x_{k+1} = f^k(x_k, u_k), \tag{19}$$

sendo x_{k+1} o próximo estado do sistema que depende do estado e da lei de controle anterior, isto é, não depende de um modelo definido, o sobrescrito k em f indica que o sistema pode ser variante no tempo.
Este trabalho aplica o princípio da otimalidade de *Bellman* para selecionar uma lei de controle u_k que minimize o índice de desempenho do sistema, que é determinada pela política de controle. Neste caso é aplicado o seguidor linear quadrático discreto com ação integral (DLQT-I) e sem ação integral (DLQT).
O valor ótimo é calculado maximizando a recompensa da Equação (17) de otimalidade de *Bellman*, a política ótima é determinada por

$$u_k = g(x_k). \tag{20}$$

O seguidor ótimo é definido pelo mapeamento g^* que promove, a maior possível, coleção de recompensas.

3.4 Sintonia do controlador LQTi via ADHDP

Neste trabalho é utilizado um método prático para solução em tempo real (*on line*) das equações, baseada na aquisição de sinais ao longo da trajetória do sistema, ao passo que não exige conhecimento da dinâmica do sistema, isto é, uma solução que não depende das matrizes A e B da representação em espaço de estados. O projeto de controladores adaptativos que aprendem em tempo real são baseados em aprendizado por reforço, conforme Lewis et al. (2012).
A Figura 2 exibe a estrutura de um algoritmo de aprendizado por reforço com ator e crítico, sendo o ator o agente responsável por gerar a política de controle, o componente crítico avalia o valor ou utilidade da ação do ator, através da equação de *Bellman*, e fornece essa informação ao ator. Essa estrutura oferece um algoritmo de avanço no tempo para determinar decisões ótimas em tempo real, conforme Lewis et al. (2012), Lewis and Vrabie (2009).
Nesta conjuntura a programação dinâmica aproximada utiliza o método *aprendizado* $-$ Q, sendo a função Q dependente da entrada de controle, isto é, dependente da ação, também chamado de ADHDP, que aproxima a função valorada de um alvo (Lewis and Vrabie, 2009). Considerando a equação do modelo da recompensa de longo termo, na forma de equação de *Bellman* (17), calcula o uso de qualquer política admissível, ótima ou não, dada por

$$Q_L(x, u) = r(x, y) + Q_L(f(x, u), g(f(x, u))), \quad \forall g, x, \tag{21}$$

sendo Q_L a função aproximadora.

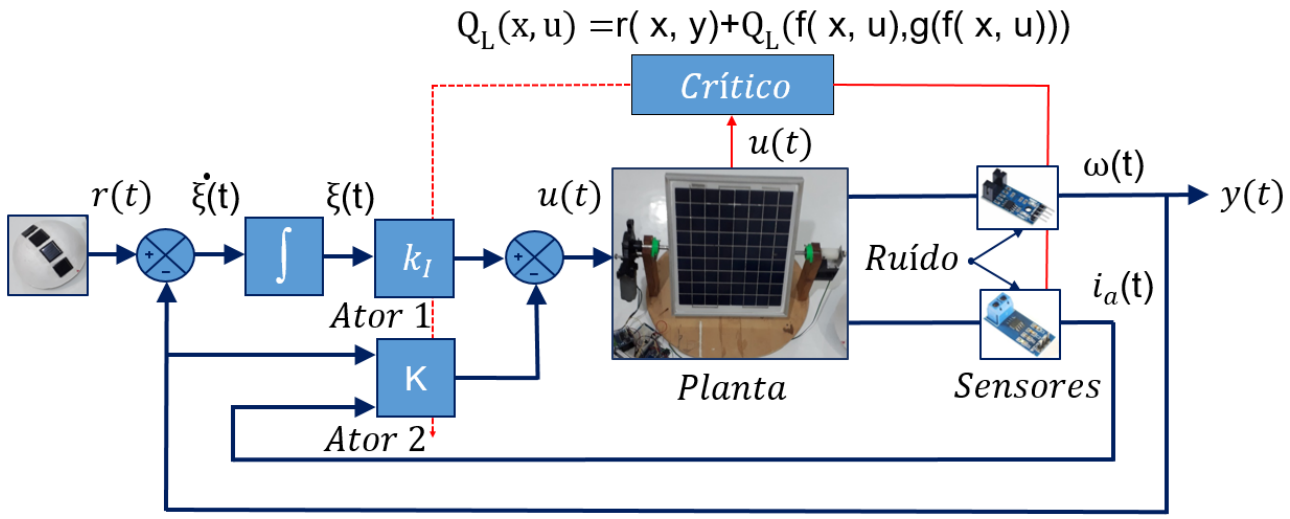

Figura 2. Diagrama de blocos do sistema seguidor com ação integral.

3.5 Algoritmo aproximado da ADHDP para sintonia do controlador

O sistema exibido na Figura 1 é um sistema discreto, descrito na forma de espaços de estados linear invariante no tempo, como mostrado nas Equações (6) e (7) e representado por

$$f(g(x)) = Ax(t) + Bu(t), \tag{22}$$

com vetor de estados $x(t)$ e entrada de controle $u(t)$. A política de controle de interesse são *feedback* das variáveis de estado, na forma

$$g(x) = K(H)x(t). \tag{23}$$

O *feedback* das variáveis de estado $x_2(t)$ e $x_3(t)$ no sistema proposto, em uma aplicação prática, pode ser aferido através do sensor de posição e velocidade instalado no eixo do módulo fotovoltaico, o estado $x_1(t)$ pode ser aferido por meio do sensor de corrente ACS712.
Para projeto do DLQT a função Q_L é dada por

$$Q_L(x_k, u_k) = x_k^T Q x_k + u_k^T R u_k + x_{k+1}^T P x_{k+1}, \tag{24}$$

sendo Q e R as matrizes de ponderação do sistema e P é a solução da equação de *Riccati* para a política prescrita K. Substituindo x_{k+1} por $Ax_k + Bu_k$ na Equação (24) e organizando na forma matricial, segundo Lewis and Vrabie (2009), resulta

$$Q_L(x_k, u_k) = \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q + A^T PA & A^T PB \\ B^T PA & R + B^T PB \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} = z_k^T H z_k. \tag{25}$$

A matriz H é obtida por meio da relação dada por

$$vec(H) = inv(Z_{bar}^T) Q_L, \tag{26}$$

sendo Z_{bar}^T a matriz transposta do produto de *kronecker* dos estados e da entrada aplicada ao sistema.
A etapa de atualização da política ótima é baseada em

$$\frac{\partial Q_L}{\partial u} = 0, \tag{27}$$

para o caso do DLQR, é definido por

$$\frac{\partial Q_L}{\partial u} = \frac{\partial}{\partial u} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} H_{xx} & H_{xu} \\ H_{ux} & H_{uu} \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \tag{28}$$

resultando na política ótima

$$u_k = -(H_{uu})^{-1} H_{ux} x_k. \tag{29}$$

O aprendizado com professor é realizado por um esquema iterativo que pode aproximar V através da aquisição dos estados da planta. A forma final para solução do controle ótimo é dada por

$$\bar{z}v(P_{i+1}) = Q_L(z, r, g(f), P_j). \tag{30}$$

A Tabela 1 exibe o algoritmo que foi utilizado para o desenvolvimento deste trabalho.

Tabela 1. Algoritmo utilizado para solução do DLQT *online*.

Programação Dinâmica - DLQT - ADHDP
Condições iniciais
Ponderação e dinâmica do sistema
$[Q, R, A_d, B_d, N] \leftarrow [\quad]$ *as matrizes* A_d *e* B_d *simulam os estados do sistema*
$H \leftarrow [\quad]$
$zbar_{(:,:)} \leftarrow [\quad], d_{alvo} \leftarrow [\quad]$
K valores iniciais
Processo iterativo
para $j \leftarrow 1 : N$
Implementação de ruído no controle
$e_i(:,i) \leftarrow [\quad]$
Ação de controle
$u_{(:,i)} \leftarrow -K_{ADHDP}x + K_{ADHDP}(I)\theta_{ref} + e_{i(:,i)}$
Grupo Z de estados e ações de controle
$Z_{(:,i)} \leftarrow [x_{(:,i)}; u_{(:,i)}]$
Atualização dos estados
$x_{j+1} \leftarrow A_d x_j + B_d u_j$
Atualização dos alvos
$d_{alvo} \leftarrow x^T Q x + u^T R u + [x_{(:,i+1)}; K_{ADHDP} x_{(:,i+1)}]^T$
$H_{ADHDP}[x_{(:,i+1)}; K_{ADHDP} x_{(:,i+1)}]$
Produto de Kronecker
$\bar{z} \leftarrow [x_{1i}^2; x_{1i}x_{2i}; \ldots; u_{1i}^2]$
se $j == n_p$
então
Atualização da matriz H por meio da Equação (26)
$H_{ADHDP} \leftarrow (\bar{Z}^T)^{-1} d_{alvo}$
Ganho ótimo de realimentação K
$H_{ux} \leftarrow H_{ADHDP(:,:)}$
$H_{uu} \leftarrow H_{ADHDP(:,:)}$
$K_{ADHDP} \leftarrow -(H_{uu})^{-1} H_{ux}$
Fim do processo iterativo

4. SIMULAÇÃO BASEADA EM MODELO

O controle ótimo de um motor CC utilizado para movimentar um painel fotovoltaico na direção do ângulo azimute para rastreamento do ponto de máxima potência via ADHDP é simulado. O algoritmo proposto é comparado com o DLQT clássico. Nesta aplicação foi adicionado um distúrbio aleatório ao sistema descrito na Equação (6), o sistema com o distúrbio pode ser expresso por

$$\dot{x}(t) = Ax(t) + Bu(t) + Ed(t), \tag{31}$$

sendo $d(t)$ o distúrbio da planta e E a matriz de ponderação do distúrbio.

4.1 Modelo dinâmico do sistema

O controle de posição do motor CC é representado por um sistema de terceira ordem, discretizado pelo método do segurador de ordem zero e com tempo de amostragem igual a 1 ms, as matrizes discretas (A_d e B_d) da Equação (6), utilizadas para simular a planta são dadas por

$$A_d = \begin{bmatrix} 0,8073 & -0,2247 & 0 \\ 0,0069 & 0,9953 & 0 \\ 0 & 0,001 & 1 \end{bmatrix}, \tag{32}$$

e

$$B_d = \begin{bmatrix} 0,2084 \\ 0,0008 \\ 0 \end{bmatrix}. \tag{33}$$

O controle de velocidade do motor CC é representado por um sistema de segunda ordem, discretizado pelo método do segurador de ordem zero e com tempo de amostragem igual a 1 ms, as matrizes discretas da Equação (8), utilizadas para simular a planta são

$$A_d = \begin{bmatrix} 0,8073 & -0,2247 \\ 0,0069 & 0,9953 \end{bmatrix}, \tag{34}$$

e

$$B_d = \begin{bmatrix} 0,2084 \\ 0,0008 \end{bmatrix}. \tag{35}$$

Devido as características do sistema as matrizes utilizadas para controle de velocidade com ação integral coincidem com as matrizes utilizadas no controle de posição do motor CC (ver Equação (15)).

4.2 Resultados

A resposta do sistema para rastreamento de posição do painel é exibida na Figura (3), considerando duas situações, a primeira sem variação paramétrica e a segunda com variação paramétrica. A referência é determinada pelo *smart wireless sensor* (apresentado na Figura 1).

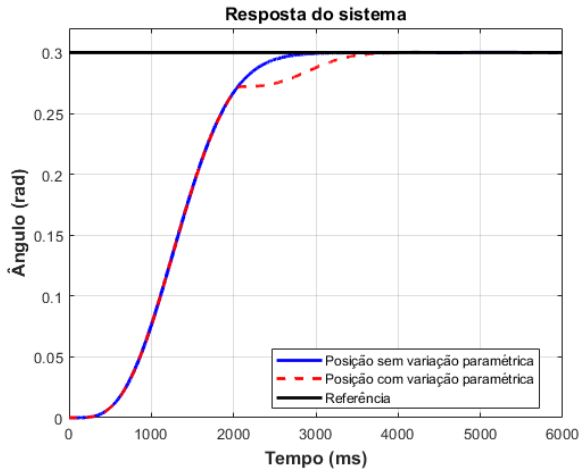

Figura 3. Resposta do sistema de rastreamento da posição do painel fotovoltaico, via ADHDP.

Conforme exibe a Figura 3, na primeira situação a posição do painel alcança a referência em aproximadamente 3 (três) segundos. A resposta do sistema, quando há variação paramétrica em dois segundos, apresenta um retardo na resposta, no entanto, continua rastreando a referência. Esta análise é importante porque alguns parâmetros do processo variam, como R_a que varia com a temperatura, ou alguma pertubação externa, como o vento, que altera a dinâmica do sistema.
O algoritmo para rastreamento da posição do painel fotovoltaico utiliza 600 iterações e apresenta como solução os ganhos ótimos

$$K_{ADHDP} = [0,3935 \;\; 0,8369 \;\; 9,4765]. \tag{36}$$

O método de Schur, apresentado em Laub (1979), resulta

$$K_{Schur} = [0,3934 \;\; 0,8267 \;\; 9,4834], \tag{37}$$

com erro normalizado máximo de 1,23%. Validando os resultados do algoritmo simulado.
A evolução da convergência dos ganhos ótimos, via ADHDP, para as duas situações supracitadas do rastreamento da posição do painel fotovoltaico é exibida na Figura 4.

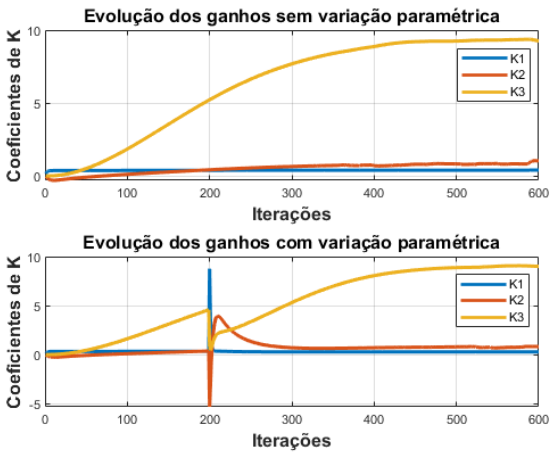

Figura 4. Evolução dos ganhos, via ADHDP, do sistema de rastreamento da posição do painel fotovoltaico.

Os ganhos k_1 e k_2, em ambas situações, convergem em um número menor de iterações em relação ao ganho k_3. No instante da variação paramétrica há mudança brusca nos ganhos, porém são estabilizados, conforme pode ser observado na Figura 4. O esforço de controle aplicado ao atuador (motor CC) para rastreamento da posição do painel fotovoltaico é exibido na Figura 5.

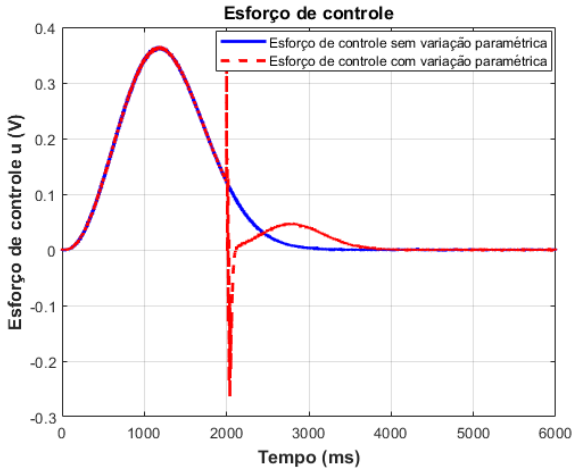

Figura 5. Esforço de controle aplicado ao sistema de rastreamento da posição do painel fotovoltaico, via ADHDP.

É importante enfatizar que em regime estacionário o esforço de controle é nulo, pois após o atuador posicionar o módulo na referência não é necessário manter a tensão aplicada ao motor. O sinal do esforço de controle para ambas situações varia de 0 a 0.36 Volts no intervalo de 0 a 1,2 s. Os sinais permanecem sobrepostos até a variação paramétrica em 2 s, conforme a Figura 5 .
A análise do rastreamento de velocidade para motores CC é necessário em aplicações que exigem velocidade fixa. A Figura 6 exibe a resposta do sistema, considerando o LQT sem ação integral, conforme descrito na Equação (11), e com ação integral.

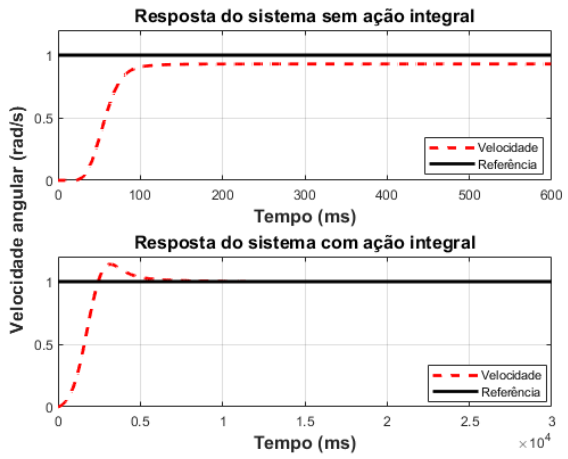

Figura 6. Resposta do sistema de rastreamento de velocidade do motor CC, via ADHDP.

A resposta do sistema sem ação integral apresenta erro de regime permanente, no entanto, esta planta tem função de transferência explicita na Equação (12) que necessita de um integrador para eliminar o erro de regime permanente com uma nova lei de controle, conforme a Equação (13). Neste caso a saída apresenta um sobressinal e não possui erro de regime permanente.
O algoritmo para rastreamento da velocidade do motor CC com ação integral apresenta os ganhos ótimos

$$K_{ADHDP} = [0,9239 \;\; 26,3451 \;\; -27,9123]\,. \tag{38}$$

O método de Schur, apresentado em Laub (1979), resulta na solução

$$K_{Schur} = [0,9239 \;\; 26,3462 \;\; -28,0543]\,, \tag{39}$$

com erro normalizado máximo de 0,5%. Validando os resultados do algoritmo simulado. A evolução da convergência para ganhos ótimos do rastreamento de velocidade do motor CC com ação integral é exibida na Figura 7.

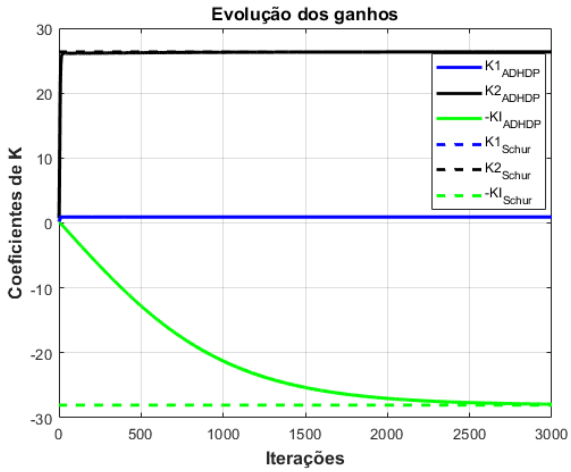

Figura 7. Resposta do sistema de rastreamento de velocidade do motor CC com ação integral.

Os ganhos k_1 e k_2 convergem rapidamente e o ganho K_I converge em aproximadamente 3000 iterações. A Figura 8 exibe o esforço de controle aplicado ao sistema com ação integral.

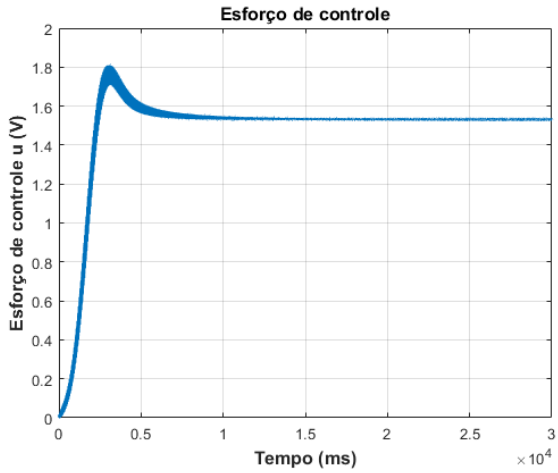

Figura 8. Resposta do sistema com ação integral.

O esforço de controle com ação integral apresenta um sobressinal. A introdução da ação integral no algoritmo proposto é essencial, pois elimina o erro de regime permanente. O algoritmo utiliza um número maior de iterações para a convergência em virtude do aumento da ordem da matriz H, em consequência da nova variável de estado ξ .

5. CONCLUSÃO

Neste trabalho foi apresentada a união entre a ADHDP e método de controle ótimo, na forma do DLQT, favorecendo o esforço de controle para três situações, sem ação integral quando a planta é do tipo 1, sem ação integral quando a planta é do tipo 0 e com ação integral para o controle de um motor CC que movimenta um painel fotovoltaico. Em comparação ao DLQT clássico a ADHDP se mostrou uma opção factível para implementação das heurísticas para sintonia QR de sistemas de controle ótimo. Observou-se que a estratégia para sintonia dos ganhos converge, mesmo quando há variação dos parâmetros.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) e Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil (CNPq). Agradeço o suporte ao grupo de trabalho do Laboratório de Sistemas Embarcados e Controle Inteligente (LABSECI) da Universidade Federal do Maranhão (UFMA) e agradeço ao Instituto Federal do Piauí (IFPI) a oportunidade de dedicação exclusiva a pesquisa.

REFERÊNCIAS

Akbar Gumilang, Y.S., Rusli, M., and Siswoyo, B. (2021). Linear quadratic regulator and luenberger observer for solar tracking system. In *2021 International Conference on Electrical and Information Technology (IEIT)*, 106–113. doi:10.1109/IEIT53149.2021.9587439.

Bai, Y., Li, R., Lin, X., Gao, S., and Jiang, Y. (2023). Research on mppt control method of high voltage bus of sar satellites. In *2023 3rd International Conference on Electrical Engineering and Control Science (IC2ECS)*, 416–421. doi:10.1109/IC2ECS60824.2023.10493439.

da Silva, Y.F., de Paiva, A.C., and da Fonseca Neto, J.V. (2021). Smart wireless sensor baseado na norma ieee 1451 para rastreamento e estimação de irradiância solar. In *Simpósio Brasileiro de Automação Inteligente-SBAI*, volume 1.

de Moura, J.P., Rego, P.H.M., and da Fonseca Neto, J.V. (2019). Online discrete-time lqr controller design with integral action for bulk bucket wheel reclaimer operational processes via action-dependent heuristic dynamic programming. *ISA transactions*, 90, 294–310.

Di Ruscio, D.L. (2012). Discrete lq optimal control with integral action: A simple controller on incremental form for mimo systems. *Modeling, Identification and Control*, 33(2), 35–44.

Gupta, S. and Biswas, P.K. (2023). State space formation and fuzzy logic control based performance analysis of dc motor drives. In *2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS)*, 1–5. doi:10.1109/CCPIS59145.2023.10291947.

Laub, A. (1979). A schur method for solving algebraic riccati equations. *IEEE Transactions on automatic control*, 24(6), 913–921.

Lewis, F.L. and Vrabie, D. (2009). Reinforcement learning and adaptive dynamic programming for feedback control. *IEEE Circuits and Systems Magazine*, 9(3), 32–50. doi:10.1109/MCAS.2009.933854.

Lewis, F.L., Vrabie, D., and Syrmos, V.L. (2012). *Optimal control*. John Wiley & Sons.

Mazumdar, D., Sinha, D., Panja, S., and Dhak, D.K. (2015). Design of lqr controller for solar tracking system. In *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1–5. doi:10.1109/ICECCT.2015.7226000.

Ogata, K. (2011). *Engenharia de controle moderno. 5ª*. São Paulo: Pearson.

Pataro, I.M.L., Gil, J.D., Guzmán, J.L., Berenguel, M., and Lemos, J.M. (2023). Optimal control of solar collector fields based on linear quadratic controller with accessible disturbance*. In *2023 European Control Conference (ECC)*, 1–6. doi:10.23919/ECC57647.2023.10178132.

Rahul, I. and Hariharan, R. (2024). Enhancement of solar pv panel efficiency using double integral sliding mode mppt control. *Tsinghua Science and Technology*, 29(1), 271–283. doi:10.26599/TST.2023.9010030.

Sen, P.C. (2021). *Principles of Electric Machines and Power Electronics, International Adaptation*. John Wiley & Sons.

Sterling, G. and Tyler, B. (2018). Renewable energy management using action dependent heuristic dynamic programming. In *2018 IEEE International Smart Cities Conference (ISC2)*, 1–5. doi:10.1109/ISC2.2018.8656942.