Optimizing Walking Controllers for Uncertain Inputs and Environments

서울대학교
박황필
Introduction

Walking is easy!
Sources of Uncertainty

• External forces
  – Wind on a gusty day
  – Pushes suddenly

• User inputs

• Transition between controllers
  – Transition from high to low speeds

• Motor noise
Goal

• Optimize 3D locomotion controllers under uncertainty
  – Find optimized control parameters

• Optimize expected objective function instead of objective function
  – Assume that all unknown quantities are modeled by probability distribution
**Controllers**

<table>
<thead>
<tr>
<th>State #</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left Foot</strong></td>
<td>Swing</td>
<td>Swing</td>
<td>Contact</td>
<td>Stance</td>
</tr>
<tr>
<td><strong>Right Foot</strong></td>
<td>Stance</td>
<td>Heel-off</td>
<td>Swing</td>
<td>Swing</td>
</tr>
</tbody>
</table>

\[ \tau = k_p (\theta_d - \theta) + k_d \dot{\theta} \]

- Control parameters
  - PD gains, target angles, start states
Objective function

- Objective (return) function $R(s_{1:T})$
- $R(s_{1:T}) = -\sum_t \sum_i \gamma_i E_i$

$$E_{\text{power}} = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{m} \tau_{t,j}^2$$

$$E_{\text{ang}} = Q(L_x; 0.04) + Q(L_y; 0.05) + Q(L_z; 0.01)$$

$$E_{\text{user}} = Q(v_x - \hat{v}_x; 0.05) + Q(s - \hat{s}; 0.05)$$

$$E_{\text{fail}} = \text{failed}_t$$
Deterministic Cases

- $s_{t+1} = f(s_t, \tau_t, e_t)$
  - $\tau_t = \pi(s_t, c_t; w) = k_p(\theta_d - \theta) + k_d \dot{\theta}$
  - $s_t$: states
  - $\tau_t$: joint torques
  - $e_t$: external forces
  - $c_t$: user inputs
  - $w$: control parameters

- Finding $w$ which maximize $R(s_{1:T})$

- But, $e_t, c_t, s_1$ are usually unknown
Random Environments

- External forces
  \[ e_t \sim p(e) \]

- User inputs
  \[ c_t \sim p(c) \]

- A start state of the controller
  \[ s_1 \sim p(s_1) \]

- Motor noise
  \[ \tau_t \sim p(\tau|s_t, c_t, w) \]
By sampling $e_t$, $c_t$, $s_1$ and $\tau_t$, we can sample an animation sequence

$$
\tau_t \sim p(\tau | s_t, c_t, w)
$$

$$
s_{t+1} = f(s_t, \tau_t, e_t)
$$

$$
s_{1:T}^{(i)} \sim p(s_{1:T} | w)
$$
Expected Return

• Optimizing *expected return* $V(w)$ instead of $R(s_{1:T})$

\[
V(w) \equiv E_{p(s_{1:T} | w)}[R(s_{1:T})] \\
= \int p(s_{1:T} | w) R(s_{1:T}) \, ds_{1:T}
\]
Optimization

• Since $V(w)$ cannot be computed analytically, we use $\hat{V}(w)$

$$\hat{V}(w) = \frac{1}{N} \sum_{i=1}^{N} R(s_{1:T}^{(i)})$$

where $s_{1:T}^{(i)} \sim p(s_{1:T} \mid w)$

• To optimize $\hat{V}(w)$, we use the CMA (Covariance Matrix Adaptation)
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate

$m_i, C_i$
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate
Covariance Matrix Adaptation

- Sample
- Select elites
- **Update mean**
- Update covariance
- Iterate
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate
Covariance Matrix Adaptation

- Sample
- Select elites
- Update mean
- Update covariance
- Iterate

$m_{i+1}, C_{i+1}$
Optimizing Walking Controllers for Uncertain Inputs and Environments

Jack M. Wang
David J. Fleet
Aaron Hertzmann

University of Toronto
Result

The graph shows the success rate (%) as a function of force magnitude (newtons) for different force directions and magnitudes.

<table>
<thead>
<tr>
<th>direction</th>
<th>0 N (baseline)</th>
<th>100 N</th>
<th>200 N</th>
<th>300 N</th>
<th>350 N</th>
<th>400 N</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 0)</td>
<td>125</td>
<td>175</td>
<td>275</td>
<td>375</td>
<td>350</td>
<td>425</td>
</tr>
<tr>
<td>(1, ±1)</td>
<td>75</td>
<td>125</td>
<td>225</td>
<td>275</td>
<td>300</td>
<td>325</td>
</tr>
<tr>
<td>(0, ±1)</td>
<td>75</td>
<td>125</td>
<td>325</td>
<td>375</td>
<td>425</td>
<td>475</td>
</tr>
<tr>
<td>(−1, ±1)</td>
<td>25</td>
<td>100</td>
<td>175</td>
<td>225</td>
<td>250</td>
<td>275</td>
</tr>
<tr>
<td>(−1, 0)</td>
<td>75</td>
<td>200</td>
<td>350</td>
<td>325</td>
<td>375</td>
<td>375</td>
</tr>
</tbody>
</table>
Result