Auto-tuning of $r$

We describe a simple heuristic to auto-tune the step size ratio $r$ on the fly. Let $\Sigma_G$ and $\Sigma_L$ be the covariance matrices for $x_G$ and $x_L$ respectively, then their minimal eigenvalues $\lambda_{\text{min},G}$ and $\lambda_{\text{min},L}$ describe the variance magnitude in the most constrained direction. Intuitively, for both HMC and Zigzag-HMC, the step size should be proportional to the diameter of this most constrained density region, which is $\sqrt{\lambda_{\text{min},G}}$ or $\sqrt{\lambda_{\text{min},L}}$. Therefore we propose a choice of $r = \sqrt{\lambda_{\text{min},L}}/\sqrt{\lambda_{\text{min},G}}$, assuming the two types of momenta lead to similar travel distance during one unit time. It is straightforward to check this assumption. At stationarity, HMC has a velocity $v_G \sim \mathcal{N}(0, I)$, so its velocity along any unit vector $u$ would be distributed as $\langle v_G, u \rangle \sim \mathcal{N}(0,1)$, and the travel distance $E|\langle v_G, u \rangle| = \sqrt{2/\pi}$. For Zigzag-HMC, as $\langle v_L, u \rangle$ does not follow a simple distribution, we estimate $E|\langle v_L, u \rangle|$ by Monte Carlo simulation and it turns out to be $\approx 0.8$, close to $\sqrt{2/\pi}$.

We test this intuitive choice of $r$ on a subset of the HIV data in [1] with 535 taxa, 5 binary and 3 continuous traits. We calculate the optimal $r = \sqrt{\lambda_{\text{min},L}}/\sqrt{\lambda_{\text{min},G}} \approx 2.5$ with $\Sigma_G$ and $\Sigma_L$ estimated from the MCMC samples. Clearly, $r$ has a significant impact on the efficiency as a very small or large $r$ leads to lower ESS (Table 1). Also, an $r$ in the order of our optimal value generates the best result, so we recommend this on-the-fly automatic tuning $r = \sqrt{\lambda_{\text{min},L}}/\sqrt{\lambda_{\text{min},G}}$ (Table 1).
Table 1: Minimal effective sample size (ESS) per running hour (hr) for partial correlation matrix elements $r_{ij}$ with different $r$ ($N = 535, P_{\text{disc}} = 5, P_{\text{cont}} = 3$). ESS values report medians across 3 independent simulations.

<table>
<thead>
<tr>
<th>$r$</th>
<th>ESS/hr min</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>32</td>
<td>266</td>
</tr>
<tr>
<td>1</td>
<td>106</td>
<td>771</td>
</tr>
<tr>
<td>10</td>
<td>118</td>
<td>855</td>
</tr>
<tr>
<td>100</td>
<td>25</td>
<td>110</td>
</tr>
</tbody>
</table>

References