Exercise 2: Sampling Methods

Versuchsdatum: 10.06.2025

TECHNISCHE UNIVERSITÄT DARMSTADT

Datenbasierte Modellierung
Maschinelles Lernen
ETIT

1 Solutions

1.1 Problem 1

1.1.1 Item (i)

The proposal distribution $q(x' \mid x)$ used in the Metropolis-Hastings algorithm is a two-dimensional Gaussian distribution centered at the current state x, with a diagonal covariance matrix $\sigma^2 I$. The figure below illustrates the shape of this distribution. It is clearly symmetric, i.e., $q(x' \mid x) = q(x \mid x')$, which simplifies the acceptance ratio in the Metropolis-Hastings update rule.

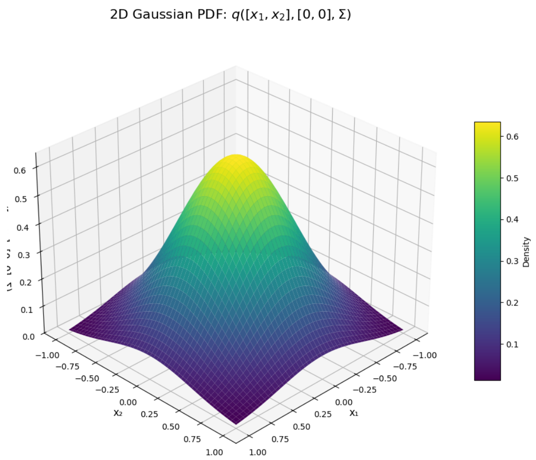

Figure 1.1: Proposal function

Before running the Metropolis-Hastings sampler, we plotted the probability density function (PDF) of the target distribution $f(x_1, x_2)$ to gain an intuition of how the generated samples should behave. From the visualization, it is clear that regions of higher probability density concentrate around the curve $x_2 = x_1^2$.

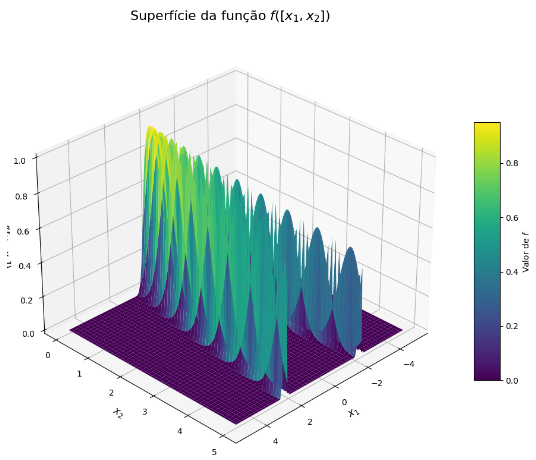

Figure 1.2: Rosenbrock plot

We then implemented the Metropolis-Hastings algorithm and plotted the generated samples in a two-dimensional plane. As expected, the sampled points concentrate around the curve $x_2 = x_1^2$, which corresponds to the regions of higher probability density in the target distribution. The resulting sequence is shown below:

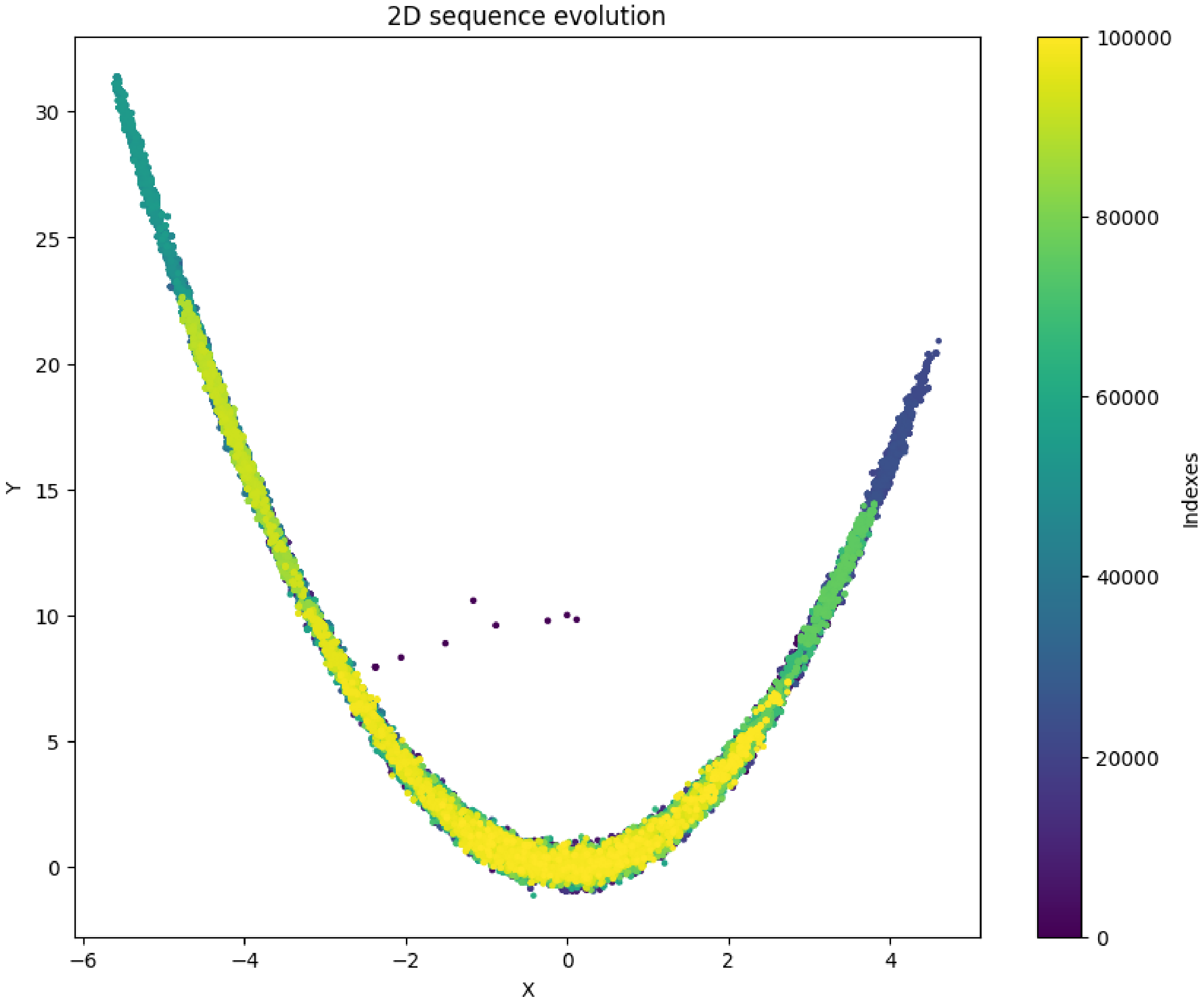

Figure 1.3: 2D sequence evolution

1.1.2 Item (ii)

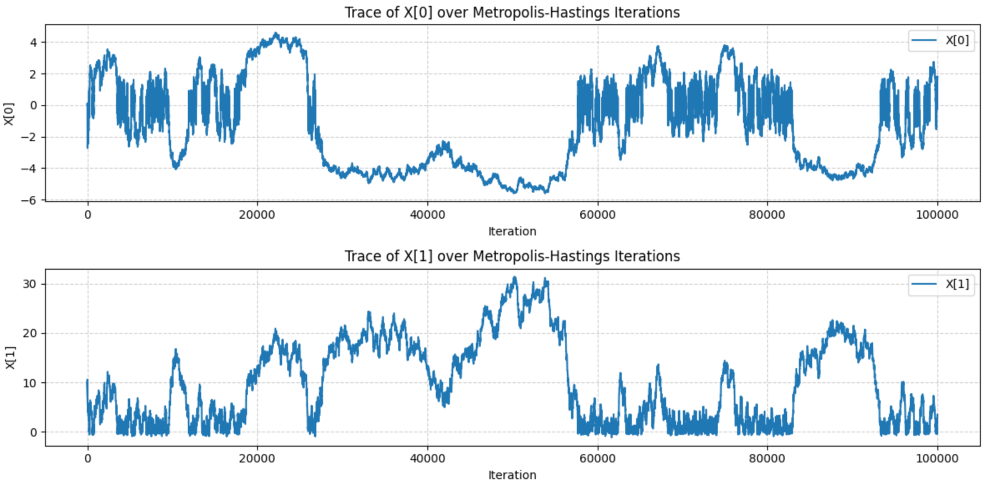

Figure 1.4: Coordinates evolution

We plotted the trace of each coordinate x_0 and x_1 over the iterations of the Metropolis-Hastings algorithm. The plots below show how the chain explores the state space throughout the sampling process. We observe that both coordinates exhibit transitions between different regions of the target distribution, as expected. The chain appears to mix well and is able to move across the high-probability regions around the curve $x_2 = x_1^2$, although some periods of slower movement and local trapping are visible.

1.1.3 item (iii)

We also analyzed how the acceptance rate of the Metropolis-Hastings algorithm varies as a function of the proposal variance σ^2. The plot below shows the acceptance rate as σ^2 increases. While we would expect the acceptance rate to decrease monotonically as σ^2 grows, the curve exhibits some fluctuations and is not perfectly monotonic. Nevertheless, a clear overall negative correlation between σ^2 and the acceptance rate can be observed.

This behavior is explained by the nature of the proposal distribution: as σ^2 increases, proposed samples tend to be further away from the current state. Since the target distribution $f(x_1, x_2)$ has a narrow high-density region around $x_2 = x_1^2$, large steps are more likely to land in low-probability regions, reducing the acceptance probability. In contrast, smaller σ^2 values lead to more local exploration and higher acceptance rates. The non-monotonicity in the curve may result from stochastic effects and the complex geometry of the target distribution.

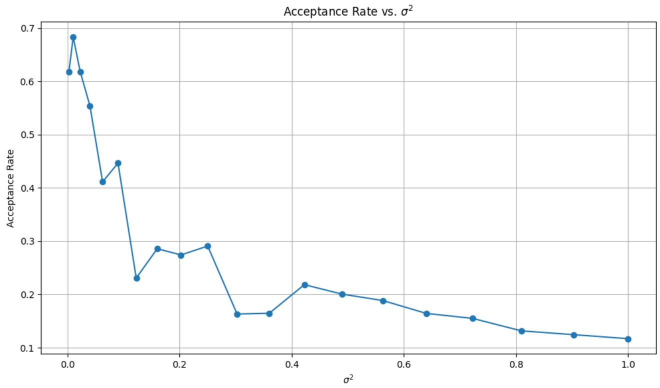

Figure 1.5: Acceptance rates

σ	Acceptance Rate
0.05	0.6185
0.10	0.6833
0.15	0.6184
0.20	0.5536
0.25	0.4114
0.30	0.4464
0.35	0.2307
0.40	0.2858
0.45	0.2741
0.50	0.2908
0.55	0.1631
0.60	0.1647
0.65	0.2183
0.70	0.2005
0.75	0.1885
0.80	0.1643
0.85	0.1550
0.90	0.1315
0.95	0.1245
1.00	0.1169

Table 1.1: Acceptance rate of Metropolis-Hastings for different values of σ.

1.1.4 item (b)

We further evaluated the efficiency of the Metropolis-Hastings sampler by analyzing the autocorrelation of the generated samples for different values of σ. The plots below show the autocorrelation function r_k as a function of lag k, computed separately for x_1 and x_2, and for three different values of σ.

As expected, increasing σ generally leads to lower autocorrelation between consecutive samples. This occurs because larger proposal steps introduce greater variation between successive points, helping the chain to explore the state space more efficiently. In contrast, when σ is small, the sampler performs only minor local moves, which results in highly correlated samples and slower mixing.

Interestingly, we also observed some unexpected behaviors. In certain cases, such as with $\sigma = 1$, the autocorrelation remains very high despite the larger proposal variance. This effect is likely due to the very low acceptance rate observed for high σ values: when the acceptance probability is low, the chain tends to reject many proposed moves and remains stuck in its current state for multiple iterations. Consequently, this leads to strongly correlated samples even though the theoretical step size is large.

Overall, these results highlight the tradeoff in selecting σ: larger values improve exploration but reduce acceptance, while smaller values increase acceptance but can result in highly autocorrelated samples.

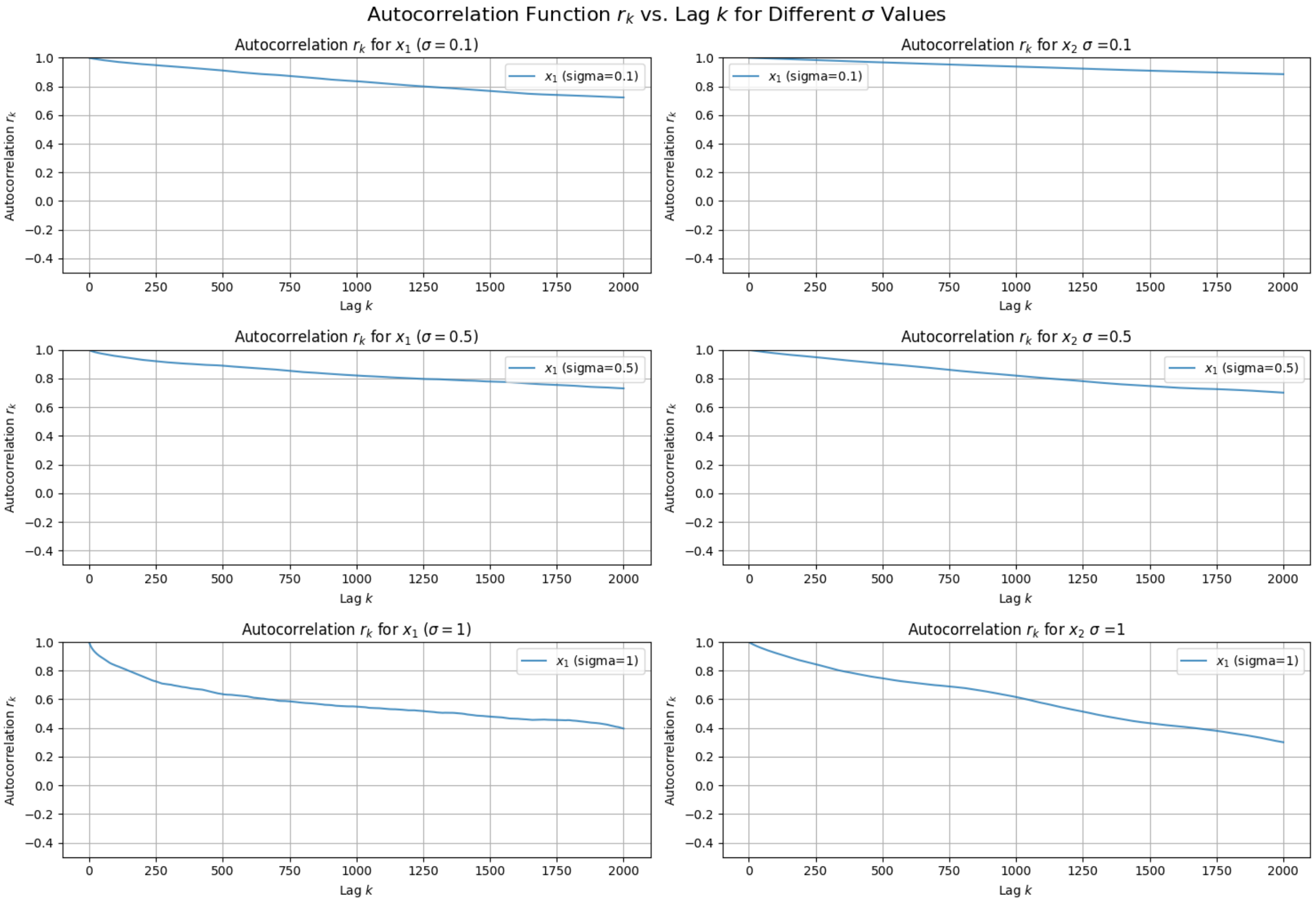

Figure 1.6: Autocorrelation graphs