High-Dimensional Gaussian Sampling

Yuejia Zhang

April 7, 2023

Problem Definition

Sampling from a d-dimensional Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where d may be large.

$$\pi(\boldsymbol{\theta}) = \frac{1}{(2\pi)^{d/2}\det(\boldsymbol{\Sigma})^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu})\right).$$

Covariance matrix $\boldsymbol{\Sigma}$ positive definite. Precision matrix $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$ exists and also positive definite.

Special Cases

- $d = 1$

Algorithm 1 Box–Muller sampler

1: Draw $u_1,\ u_2 \overset{\text{i.i.d.}}{\sim} \mathcal{U}((0,1])$.
2: Set $\tilde{u}_1 = \sqrt{-2\log(u_1)}$.
3: Set $\tilde{u}_2 = 2\pi u_2$.
4: **return** $(\theta_1, \theta_2) = \left(\mu + \frac{\tilde{u}_1}{\sqrt{q}} \sin(\tilde{u}_2), \mu + \frac{\tilde{u}_1}{\sqrt{q}} \cos(\tilde{u}_2)\right)$.

Special Cases

Algorithm 2 Sampler when $\mathbf{Q}$ is a diagonal matrix

1: **for** $i \in [d]$ **do** $\triangleright$ In some programming languages, this loop can be vectorized.
2: Draw $\theta_i \sim \mathcal{N}(\mu_i, 1/q_i)$.
3: **end for**
4: **return** $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_d)^\top$.

General Cases

Algorithm 3 Cholesky sampler

1: Set $\mathbf{C} = \mathrm{chol}(\mathbf{Q})$. ▷ $\mathbf{Q} = \mathbf{C}\mathbf{C}^\top$
2: Draw $\mathbf{z} \sim \mathcal{N}(\mathbf{0}_d, \mathbf{I}_d)$.
3: Solve $\mathbf{C}^\top \mathbf{w} = \mathbf{z}$ w.r.t. $\mathbf{w}$.
4: **return** $\boldsymbol{\theta} = \boldsymbol{\mu} + \mathbf{w}$.

Problem:

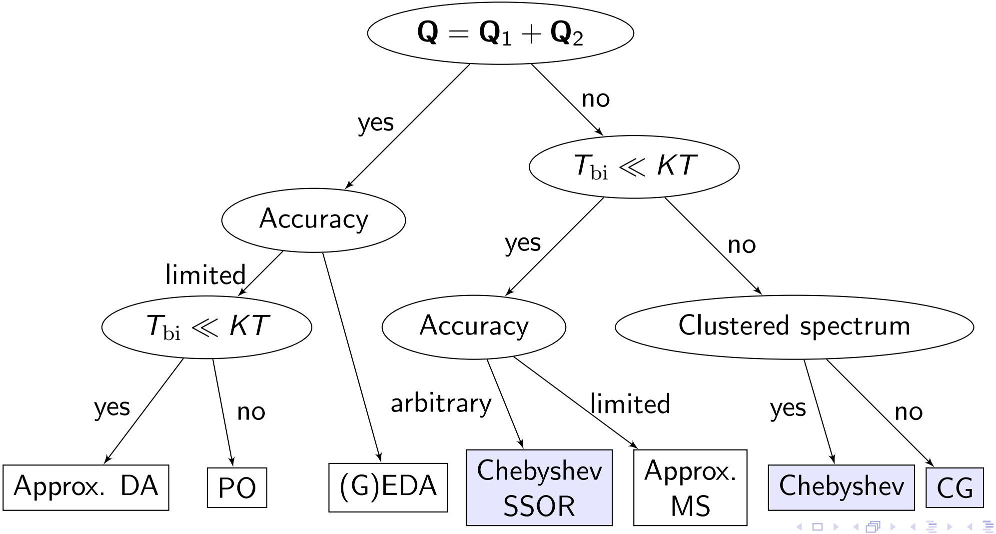
- Storage requirement $\Theta(d^2)$.
- Computational cost $\mathcal{O}(d^3 + d^2 T)$ (T is the number of samples), only when $\mathbf{Q}$ is unchanged.

More Efficient Solutions

- Square Root approximation: Approximate $\mathbf{Q}^{1/2}$.
- Conjugate Gradient: Solve a linear system w.r.t. $\mathbf{Q}$.
- Matrix Splitting: A generalization of Gibbs Sampler.
- Data Augmentation: Introduce auxiliary variable.

Improvement: computational cost $\mathcal{O}(d^2 T)$ and storage requirement $\Theta(d)$.

How to Choose the Sampler



1

1

1

1

PPA

1