Categorical Reparametrization with Gumbel-Softmax

Quiz, 9 questions

1
point

1.

What can discrete latent variables in models correspond to?

☐ Air temperature outside

☑ Image regions

☑ Semantic classes

☐ Program execution time

1
point

2.

Why can't we use discrete random variables "as is" in neural networks?

◉ Sampling from them is not differentiable

◯ They are hard to parametrise

◯ They are usually hard to sample from

◯ They exist only in low dimensions

1
point

3.

Select Gumbel(0, 1) distribution

◉ $\mathrm{Gumbel}(x;\, 0,\, 1) \propto e^{-(x+e^{-x})}, \; x \in (-\infty, \infty)$

◯ $\mathrm{Gumbel}(x;\, 0,\, 1) \propto e^{-x}, \; x \in (0, \infty)$

◯

Categorical Reparametrization with Gumbel–Softmax

---

1 point

4.

What is a simple way get sample g from $\mathrm{Gumbel}(0, 1)$

○ $u \sim U(0, 1), g \sim e^u$

○ $u \sim U(0, 1), g \sim e^{-u}$

◉ $u \sim U(0, 1), g \sim -log(-log(u))$

○ $u \sim U(0, 1), g \sim -log(u)$

---

1 point

5.

In a Gumbel-Max trick to draw samples from $\mathrm{Cat}(\pi)$, one can compute $z = \mathrm{one\_hot}(arg \max_i[g_i + \log \pi_i])$, where g_i are i.i.d. samples from Gumbel(0, 1) distrubtion. How can we approximate $\mathrm{one\_hot}(arg \max[\ldots])$ operation in a differentiable manner?

○ By directly sampling from categorical distribution

○ By sorting indices of $[g_i + \log \pi_i]$ and taking the index of a maximum value

○ By sampling from Gumbel(0, 1) distribution

◉ By using Softmax operation over $[g_i + \log \pi_i]$

---

1 point

6.

For which values of τ is it possible to compute derivatives of Gumbel Softmax distribution $\partial y / \partial \pi$?

◉ $\tau > 0$

○ $\tau \le 0$

$\tau > 0$

$\tau < 0$

| 1 |
| point |

7.

How does parameter τ influence the behaviour of Gumbel-Softmax?

☐ As τ approaches to 0, samples are far from one-hot

☑ As τ approaches to 0, samples are close to one-hot

☐ As τ approaches to 0, variance of the gradient becomes small

☑ As τ approaches to 0, variance of the gradient becomes large

---

| 1 |
| point |

8.

What is a Straight-through Gumbel-Softmax estimator?

○ z is sampled from Gumbel(0, 1), gradient is computed from discrete estimation

○ z is sampled from discrete distribution, gradient is computed from discrete estimation

◉ z is sampled from discrete distribution, gradient is computed from continuous estimation

○ z is sampled from continuous distribution, gradient is computed from discrete estimation

---

| 1 |
| point |

9.

How can one deal with unlabeled data for classification problem?

☐ Use VAE with continuous latent variables

☐ Use reparametrization trick

☑ Marginalize over all classes

Categorical Reparametrization with Gumbel–Softmax

Quiz, 9 questions

☐ I, **Jiadai Zhao**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Submit Quiz