FAKULTÄT FÜR WIRTSCHAFTSINFORMATIK
UND WIRTSCHAFTSMATHEMATIK

UNIVERSITÄT
MANNHEIM

Prof. Dr. Leif Döring
Sara Klein, Benedikt Wille

Reinforcement Learning

8. Solution Sheet

1. Convergence of Stochastic Gradient Descent

The goal of this exercise is to prove the convergence of the stochastic version of the gradient descent method. Let $F : \mathbb{R}^d \to \mathbb{R}$ be a function of the form $F(x) = \mathbb{E}[f(x, Z)]$ for some $Z \sim \mu$, whose minimum we want to find but whose gradient we cannot exactly compute. The idea is to approximate the gradient of F by $\nabla_x f(x, Z_i)$ with independent realisations $Z_i \sim \mu$ in each step, leading to the following algorithm:

Algorithm 1: Stochastic Gradient Descent

```
Data: Realisation of initial random variable $X_0$, stepsizes $\alpha_k$
Result: Approximation X of a stationary point of $F$
Set $k = 0$
while not converged do
    simulate $Z_{k+1} \sim \mu$ independently
    approximate the gradient $\nabla_x F(X_k)$ through
    $G_k = \nabla_x f(X_k, Z_{k+1})$
    set $X_{k+1} = X_k - \alpha_k G_k$
    set $k = k + 1$
end
return $X := X_k$
```

Assume the following:

- Let $(\Omega, \mathcal{F}, (\mathcal{F}_k)_{k\in\mathbb{N}}, \mathbb{P})$ be a filtered probability space, where the filtration is defined by

$$\mathcal{F}_k := \sigma(X_0, Z_m, m \leq k) \text{ for } Z_k \sim_{\text{i.i.d}} \mu,$$

- let $F : \mathbb{R}^d \to \mathbb{R}$, $x \mapsto \mathbb{E}[f(x, Z)]$ for $Z \sim \mu$ be an L-smooth function for some $L < 1$, i.e.

$$\|\nabla F(x) - \nabla F(y)\| \leq L\|x - y\| \quad \forall x, y \in \mathbb{R}^d$$

  and let $F_* := \inf_{x\in\mathbb{R}^d} F(x) > -\infty$,
- let $\nabla_x F(x) = \mathbb{E}[\nabla_x f(x, Z)]$ and $\mathbb{E}[\|\nabla_x f(x, Z)\|^2] \leq c$ for some $c > 0$ and all $x \in \mathbb{R}^d$,
- let $(\alpha_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{F}_k$-adapted and strictly positive random variables, where

$$\sum_{k=1}^{\infty} \alpha_k = \infty \text{ and } \sum_{k=1}^{\infty} \alpha_k^2 < \infty$$

- let X_0 be such that $\mathbb{E}[F(X_0)] < \infty$, and
- let $(X_k)_{k\in\mathbb{N}}$ be the random variables generated by applying Stochastic Gradient Descent.

a) For all L-smooth functions $f : \mathbb{R}^d \to \mathbb{R}$ it holds that

$$f(x+y) \leq f(x) + y^T \nabla f(x) + \frac{L}{2}\|y\|^2 \quad \forall x, y \in \mathbb{R}^d.$$

Solution:

Let $x, y \in \mathbb{R}^d$ be fixed. We define $\phi(t) := f(x+ty)$ for all $t \in [0,1]$ and apply the chain rule in order to derive

$$\phi'(t) = y^T \nabla f(x+ty) \quad \forall t \in [0,1].$$

By the fundamental theorem of calculus it follows

$$\begin{aligned} f(x+y) - f(x) = \phi(1) - \phi(0) &= \int_0^1 \phi'(t)\,dt = \int_0^1 y^T \nabla f(x+ty)\,dt \\ &= \int_0^1 y^T \nabla f(x)\,dt + \int_0^1 y^T(\nabla f(x+ty) - \nabla f(x))\,dt \\ &\leq y^t \nabla f(x) + \int_0^1 \|y\| \cdot \|\nabla f(x+ty) - \nabla f(x)\|\,dt \\ &\leq y^T \nabla f(x) + \|y\| \int_0^1 Lt \cdot \|y\|\,dt \\ &= y^T \nabla f(x) + \frac{L}{2}\|y\|^2, \end{aligned}$$

where we have applied Cauchy-Schwarz followed by the L-smoothness of f.

b) Define $M_{k+1} := \nabla_x F(X_k) - \nabla_x f(X_k, Z_{k+1})$ and show that

$$\mathbb{E}[M_{k+1}|\mathcal{F}_k] = 0 \text{ and } \mathbb{E}[\|M_{k+1}\|^2|\mathcal{F}_k] \leq c - \|\nabla_x F(X_k)\|^2 \quad \forall k \in \mathbb{N}.$$

Solution:

Since by definition of the filtration X_k is $\mathcal{F}_k$-measurable and Z_{k+1} is independent of $\mathcal{F}_k$ we can compute

$$\mathbb{E}[M_{k+1}|\mathcal{F}_k] = \nabla_x F(X_k) - \mathbb{E}[\nabla_x f(\,\cdot\,, Z_{k+1})](X_k) \overset{ass.}{=} 0$$

and

$$\begin{aligned} \mathbb{E}[\|M_{k+1}\|^2|\mathcal{F}_k] = \|\nabla_x F(X_k)\|^2 &- 2\mathbb{E}[\langle \nabla_x F(\,\cdot\,), \nabla_x f(\,\cdot\,, Z_{k+1})\rangle](X_k) \\ &+ \mathbb{E}[\|\nabla_x f(\,\cdot\,, Z_{k+1})](X_k) \\ &\overset{ass.}{\leq} c - \|\nabla_x F(X_k)\|^2. \end{aligned}$$

c) Show that $\lim_{k\to\infty} F(X_k) = F_\infty$ almost surely for some almost surely finite random variable.

Solution:

Using a) and b) we obtain (path-wise) that

$$\begin{aligned} F(X_{k+1}) &= F(X_k - \alpha_k \nabla_x f(X_k, Z_{k+1})) \\ &\leq F(X_k) - \alpha_k \langle \nabla_x F(X_k), \nabla_x f(X_k, Z_{k+1})\rangle + \alpha_k^2 \frac{L}{2}\|\nabla_x f(X_k, Z_{k+1})\|^2 \\ &= F(X_k) - \alpha_k \|\nabla_x F(X_k)\|^2 + \alpha_k \langle \nabla_x F(X_k), M_{k+1}\rangle \\ &+ \alpha_k^2 \frac{L}{2}(\|\nabla_x F(X_k)\|^2 - 2\langle \nabla_x F(X_k), M_{k+1}\rangle + \|M_{k+1}\|^2) \end{aligned}$$

and therefore, using again that X_k *is* $\mathcal{F}_k$*-measurable,*

$$\mathbb{E}[F(X_{k+1}) - F_*|\mathcal{F}_k] \leq (F(X_k) - F_*) + \alpha_k^2 \frac{L}{2} c - \alpha_k \|\nabla_x F(X_k)\|^2.$$

Now a direct application of the Robbins-Siegmund Theorem 4.4.2. with $Z_k = F(X_k) - F_*$*,* $A_k = 0$*,* $B_k = \alpha_k^2 \frac{L}{2} c$*, and* $C_k = \alpha_k \|\nabla_x F(X_k)\|^2$ *yields the assertion. All random variables are positive because of the definition of* F_* *and the fact that all* $\alpha_k > 0$ *by assumption and the summation conditions of the theorem hold because of the assumptions on* α_k*, justifying its application.*

d) Show that $\lim_{k\to\infty} \|\nabla_x F(X_k)\|^2 = 0$ almost surely.

Solution:

The application of the Robbins-Siegmund Theorem in part c) reveals that almost surely it holds $\sum_{k=0}^{\infty} \alpha_k \|\nabla_x F(X_k)\|^2 < \infty$*. Since* $\sum_{k=0}^{\infty} \alpha_k = \infty$ *almost surely, there can not exist any* $\epsilon > 0$ *such that on a non-null set of* Ω *it holds* $\|\nabla_x F(X_k(\omega))\|^2 > \epsilon$ *for all* $k \geq \bar{k}(\omega) \geq 0$ *for some* $\bar{k}(\omega)$*. Thus almost surely*

$$\liminf_{k\to\infty} \|\nabla_x F(X_k)\| = 0.$$

Now let ω *be a path on which the sum over* $\alpha_k \|\nabla_x F(X_k)\|^2$ *is finite and the sum over* α_k *is infinite. Assume that*

$$\limsup_{k\to\infty} \|\nabla_x F(X_k(\omega))\|^2 \geq \epsilon^2 > 0$$

and consider two sub-sequences $(m_j(\omega))_{j\in\mathbb{N}}$*,* $(n_j(\omega))_{j\in\mathbb{N}}$*, with* $m_j(\omega) < n_j(\omega) < m_{j+1}(\omega)$ *such that*

$$\frac{\epsilon}{3} < \|\nabla_x f(X_k(\omega))\| \quad \text{for } m_j(\omega) \leq k < n_j(\omega)$$

and

$$\|\nabla_x f(X_k(\omega))\| \leq \frac{\epsilon}{3} \quad \text{for } n_j(\omega) \leq k < m_{j+1}(\omega).$$

Such subsequences must exist, because we proved, that the limes inferior is zero. Moreover, let $\bar{j}(\omega) \in \mathbb{N}$ *be sufficiently large such that*

$$\sum_{k=m_{\bar{j}(\omega)}}^{\infty} \alpha_k(\omega) \|\nabla_x F(X_k(\omega))\|^2 \leq \frac{\epsilon^2}{9L}.$$

Using L-smoothness for all $j \geq \bar{j}(\omega)$ and $m_j(\omega) \leq m \leq n_j(\omega) - 1$ it holds true that

$$\begin{aligned}
\mathbb{E}[\|\nabla_x F(X_{n_j(\omega)}) - \nabla_x F(X_m)\| | \mathcal{F}_m](\omega) &\leq \sum_{k=m}^{n_j(\omega)-1} \mathbb{E}[\|\nabla_x F(X_{k+1}) - \nabla_x F(X_k)\| | \mathcal{F}_k](\omega) \\
&\leq L \sum_{k=m}^{n_j(\omega)} \mathbb{E}[\|X_{k+1} - X_k\| | \mathcal{F}_k](\omega) \\
&= \sum_{k=m}^{n_j(\omega)} \alpha_k(\omega) \mathbb{E}[\|\nabla_x f(X_k, Z_{k+1})\| | \mathcal{F}_k] \\
&= \sum_{k=m}^{n_j(\omega)} \alpha_k(\omega) \|\nabla_x F(X_k(\omega)\| \\
&\leq L \frac{3}{\epsilon} \sum_{k=m}^{n_j(\omega)} \alpha_k(\omega) \|\nabla_x F(X_k(\omega)\|^2 \\
&\leq \frac{\epsilon}{3},
\end{aligned}$$

where we have used that $\|\nabla_x F(X_k)(\omega)\| > \frac{\epsilon}{3}$ for $m_j(\omega) \leq k \leq n_j(\omega) - 1$. This implies that

$$\|\nabla_x F(X_m(\omega))\| \leq \mathbb{E}[\|\nabla_x F(X_{n_j(\omega)})\| | \mathcal{F}_m](\omega) + \frac{\epsilon}{3} \leq \frac{2\epsilon}{3}$$

and therefore $\|\nabla_x F(X_m(\omega))\| \leq \frac{2\epsilon}{3}$ for all $m \geq m_j(\omega)$. This is in contradiction to

$$\limsup_{k \to \infty} \|\nabla_x F(X_k(\omega))\|^2 \geq \epsilon^2.$$

Thus, the assertion holds.