Proof of the Lemmas

Liwei Wang
Key Laboratory of Machine Perception, MOE,
School of Electronics Engineering and Computer Science, Peking University,
wanglw@cis.pku.edu.cn

Here we give the formal proofs of Lemma 8 and Lemma 9. Actually they are simple consequences of the following two theorem respectively.

Theorem 1 *Let f be a function defined on $[0,1]^d$ and is Kth order smooth. Let $r = \int_{[0,1]^d} |f(x)|dx$, then $\|f\|_\infty = O(r^{\frac{K}{K+d}}) = O(r \cdot (\frac{1}{r})^{\frac{d}{K+d}})$, where $\|f\|_\infty = \sup_{x\in[0,1]^d} |f(x)|$.*

Theorem 2 *Let f be a function defined on $[0,1]^d$ and is infinitely smooth. If $\int_{[0,1]^d} |f(x)|dx = r$, then $\|f\|_\infty = O(r \cdot \log^d(\frac{1}{r}))$.*

Proof of Lemma 8 By the assumption that $|\tilde{\Phi}(x)| \le \frac{1}{\alpha}|\Phi(x)|$ for all $x \in [0,1]^d$, we have

$$\int_{[0,1]^d} |\tilde{\Phi}(x)|dx = O(r).$$

Since $\tilde{\Phi}$ is Kth order smooth, by Theorem 1 we have

$$\|\tilde{\Phi}\|_\infty = O\left(r \cdot \left(\frac{1}{r}\right)^{\frac{d}{K+d}}\right).$$

Therefore

$$\|\Phi\|_\infty \le \beta\|\tilde{\Phi}\|_\infty = O\left(r \cdot \left(\frac{1}{r}\right)^{\frac{d}{K+d}}\right).$$

■

Lemma 9 can be proved in the same way by Theorem 2.

Below, we give the proofs of Theorem 1 and Theorem 2.

Proof of Theorem 1 We first consider the one-dimensional case, i.e. $d = 1$. Note that if $f \in F_C^K$, then

$$|f^{(K-1)}(x) - f^{(K-1)}(x')| \le C|x - x'| \tag{1}$$

for all $x, x' \in [0,1]$. Hence we relax the constraint that f is Kth order smooth to (1). The idea of the proof for $d = 1$ is that, (one of) the optimal f (i.e. $\|f\|_\infty$ achieves the maximum) under the constraint (1) is of the form

$$f(x) = \begin{cases} \frac{C}{K!}|x-\xi|^K & : \quad 0 \le x \le \xi, \\ 0 & : \quad \xi < x \le 1. \end{cases} \tag{2}$$

That is

$$|f^{(K-1)}(x) - f^{(K-1)}(x')| = C|x - x'|$$

for all $x, x' \in [0,\xi]$, where ξ is determined by $\int_0^1 f(x)dx = r$. It is then easy to check that $\|f\|_\infty = O(r^{\frac{K}{K+1}})$.

For the formal proof, assume that $f(x) \geq 0$ for all $x \in [0,1]$. Let f_0 be the optimal function, i.e. $\|f_0\|_\infty \geq \|f\|_\infty$ for all f satisfying the constraints. We will show that

$$|f_0^{(K-1)}(x) - f_0^{(K-1)}(x')| = C|x - x'|$$

for all x, x' such that $f(x) > 0$ and $f(x') > 0$. Assume, for the sake of contradiction, that this is not true. Then there exits an interval (a, b) and two constants C_1, C_2, such that

$$f_0(x) \geq C_1 > 0$$

and

$$|f_0^{(K-1)}(x) - f_0^{(K-1)}(x')| \leq C_2|x - x'| < C|x - x'|$$

for all $x, x' \in (a, b)$. Let

$$F(x) = \int_0^x f_0(t)dt.$$

We have

$$F(0) = 0, \qquad F(1) = r, \qquad F'(x) \geq 0$$

for all $x \in [0,1]$. We also have

$$F'(x) \geq C_1 > 0, \qquad |F^{(K)}(x) - F^{(K)}(x')| \leq C_2|x - x'| < C|x - x'|$$

for all $x, x' \in (a, b)$. Moreover, $\|F'\|_\infty$ achieves the maximum.

Now, we will construct a function $h(x)$, so that there is a small γ so that $F + \gamma h$ satisfies all the constraints but

$$\|F' + \gamma h'\|_\infty > \|F'\|_\infty,$$

which leads to a contradiction.

Denote $x^* = \arg\max_{x \in [0,1]} F'(x)$. We will discuss three cases:

$$x^* \in (a, b), \qquad x^* \in [b, 1], \qquad x^* \in [0, a].$$

If $x^* \in (a, b)$, let

$$h(x) = (x - a)^{K+1}(b - x)^{K+1}.$$

It is easy to check that for $|\gamma|$ sufficiently small,

$$(F + \gamma h)'(x) \geq 0, \quad x \in [0, 1],$$

and

$$|(F + \gamma h)^{(K)}(x) - (F + \gamma h)^{(K)}(x')| \leq C|x - x'|, \quad x, x' \in [0, 1]$$

If $x^* \in (a, \frac{a+b}{2})$, take $\gamma > 0$; if $x^* \in (\frac{a+b}{2}, b)$, take $\gamma < 0$. It is clear that in both cases

$$(F' + \gamma h')(x^*) > F'(x^*).$$

If $x^* = \frac{a+b}{2}$, we can just use $b' = a + \frac{3}{4}(b - a)$ instead of b.

If $x^* \in [b, 1]$, let

$$h(x) = \begin{cases} 0 & 0 \leq x \leq a, \\ \frac{(x-a)^{K+1}}{(x-a)^{K+1} + (b-x)^{K+1}} \cdot \frac{x-1}{1-b} & a < x < b, \\ \frac{x-1}{1-b} & b \leq x \leq 1. \end{cases}$$

It is not difficult to check that for sufficiently small $\gamma > 0$,

$$(F + \gamma h)'(x) \geq 0,$$

and

$$|(F + \gamma h)^K(x) - (F + \gamma h)^K(x')| \leq C|x - x'|$$

for all $x, x' \in [0, 1]$, but

$$(F + \gamma h)'(x^*) > F'(x^*).$$

The case $x^* \in [0, a]$ can be treated in the same way.

Now we have proved that the optimal f is, on the interval that $f(x) > 0$, a Kth order polynomial with the coefficient of the term x^K is $\frac{C}{K!}$. If $f(x) > 0$ only on $[0, \xi)$ ($\xi < 1$), then f must be of the form in Eq.(2). This is because f has continuous derivatives up to order $K-1$ at ξ, hence the derivatives up to $K-1$th order must vanish at ξ. Thus we only need to exclude the possibility that $f(x) > 0$ on $[0, 1]$ except at a finite number of zeros. Below we will show that this is not possible because such a f must have $\int_0^1 |f(x)|dx$ is greater than some constant, which contradicts to $\int_0^1 |f(x)|dx = r$ where r can be arbitrarily small.

Let f be represented by the following standard form

$$f(x) = \frac{C}{K!}(x-r_1)^{p_1} \dots (x-r_t)^{p_t}[(x-r_{t+1})^{2p_{t+1}} + \alpha_1] \dots [(x-r_s)^{2p_s} + \alpha_{s-t}].$$

where $\alpha_i \geq 0$ and the powers summing up to K. So the first t terms correspond to real zeros and the others correspond to complex zeros. In fact we only need to consider the case that all $\alpha_i = 0$, since it is easy to see that positive α_i increase $\int |f(x)|dx$. Therefore we assume f has only real zeros. We first assume that there is no zero in $[0, 1]$ and p is the total power of all negative zeros. Then

$$\begin{aligned}
\int_0^1 |f(x)|dx &\geq \frac{C}{K!}\int_0^1 x^p(1-x)^{K-p}dx \\
&= \frac{C}{K!}\frac{\Gamma(p+1)\Gamma(K-p+1)}{\Gamma(K+2)} \\
&\geq \frac{C}{K!}\frac{\Gamma^2(\frac{K}{2}+1)}{\Gamma(K+2)},
\end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function.

For the case that there are zeros in $[0, 1]$. Assume without loss of generality that $0 \leq r_1 < r_2 < \dots < r_l \leq 1$. Denote $\Delta_1 = r_1, \Delta_2 = r_2 - r_1, \dots, \Delta_{l+1} = 1 - r_l$. We must have $\max \Delta_i \geq \frac{1}{K+1}$. Let i^* be the corresponding i, and consider r_{i^*-1} and r_{i^*}. Then we have

$$\begin{aligned}
\int_0^1 |f(x)|dx &\geq \frac{C}{K!}\int_{r_{i^*-1}}^{r_{i^*}} (x - r_{i^*-1})^p (r_{i^*} - x)^{K-p}dx \\
&\geq \frac{C}{K!}(\frac{1}{K+1})^{K+1}\frac{\Gamma^2(\frac{K}{2}+1)}{\Gamma(K+2)}.
\end{aligned}$$

Thus in either case the integral can not be arbitrarily small.

To conclude the $d = 1$ case, the optimal function f_0 must satisfy

$$f_0^{(K-1)}(x) - f_0^{K-1}(x') = C|x - x'|$$

for all x, x' such that $f(x) > 0$ and $f(x') > 0$. Since f must have continuous derivatives up to the $(K-1)$th order, (one of) the optimal f has to be of the form given in Eq(2). This completes the proof of the one-dimensional case.

For the general case $d \geq 1$, the idea is to relax the constraints that the partial derivatives are Lipschitz to that the directional partial derivatives are Lipschitz.

First note that all $K-1$th order partial derivatives are Lipschitz implies that all the $K-1$th order directional derivatives are Lipschitz too. To be precise, let u be a unit vector, i.e. $\|u\| = 1$. Also let $\phi_{x,u}(t) = f(x + tu)$, where x is arbitrary. Then the pth order directional derivative is defined as $\phi_{x,u}^{(p)}(t)$. It is clear by calculus that if all $D^{\mathbf{k}}f$ are Lipschitz with some constant C for all $\mathbf{k}$ such that $|\mathbf{k}| = K-1$, then $\phi_{x,u}^{(K-1)}(t)$ is Lipschitz with some other constant C' for all t, x and u. Now, let $\mathbf{0}$ be the d-dimensional vector $(0, \dots, 0)$ and $x_0 \in [0, 1]^d$. Let $\phi_{x_0}(t) = f(\mathbf{0} + t\frac{x_0}{\|x_0\|})$.

According to the arguments for the one dimensional case, it is not difficult to see that if $\phi_{x_0}^{(K-1)}$ is Lipschitz for all t and $x_0 \in [0, 1]^d$ with constant C', then (one of) the optimal ϕ mush be of the form

$$\phi_{x_0}(t) = \begin{cases} \frac{C'}{K!}|t - \xi|^K & 0 \leq t \leq \xi, \\ 0 & \xi < t. \end{cases}$$

Hence the corresponding f has the form[1]

$$f(x) = \begin{cases} \frac{C'}{K!}|\|x\| - \xi|^K & 0 \leq \|x\| \leq \xi, \\ 0 & \xi < \|x\|. \end{cases}$$

where ξ is determined by $\int_{[0,1]^d} |f(x)|dx = r$. Finally, simple calculations show that $\|f\|_\infty = O(r^{\frac{K}{K+d}})$. This completes the proof. ■

Proof of Theorem 2 First consider the $d = 1$ case. Since f is infinitely smooth, it is Kth order smooth for arbitrary large K. Hence we can choose K depending on r. Let

$$K + 1 = \frac{\log \frac{1}{r}}{\log\log \frac{1}{r}}.$$

We know that the optimal f is of the form in (2). We point out that this K is (approximately) the largest K such that (2) is still the optimal form. If K is larger than this, ξ will be out of $[0, 1]$, and the argument in the proof of Theorem 1 does not hold. Since $\int_0^1 |f(x)| = r$, we have

$$\xi^{K+1} = \frac{(K+1)!}{C}.$$

It is clear that

$$\|f\|_\infty = \frac{C}{K!}\xi^K.$$

Remember that

$$K + 1 = \frac{\log \frac{1}{r}}{\log\log \frac{1}{r}},$$

also note that

$$(\frac{1}{r})^{\frac{\log\log \frac{1}{r}}{\log \frac{1}{r}}} = \log \frac{1}{r},$$

then by Stirling's formula, it is easy to show that $\|f\|_\infty = O(r \cdot \log \frac{1}{r})$.

For the general $d \geq 1$ case, take

$$K + d = \frac{\log \frac{1}{r}}{\log\log \frac{1}{r}}.$$

By similar arguments in the proof Theorem 1 we have $\|f\|_\infty = O(r \cdot \log^d \frac{1}{r})$. ■

[1] f is optimal under the relaxed constraints of directional partial derivatives. Actually this f no longer satisfies the original partial derivative constraints.