(1) Just prove that $h_\varepsilon^k(t)$ is continuous in $t=0$， $t=-\varepsilon$

when $t=0$，

$$\lim_{t\to 0^-} h_\varepsilon^k(t)=\varepsilon^{k(1-b)}b^{-k}\varepsilon^{kb}=\varepsilon^k b^{-k}\ ,\ \ \lim_{t\to 0^+} h_\varepsilon^k(t)=\lim_{t\to 0^+}(\varepsilon b^{-1})^k=\varepsilon^k b^{-k}$$

Then: $\lim_{t\to 0^-} h_\varepsilon^k(t)=\lim_{t\to 0^+} h_\varepsilon^k(t)$；

when $t=-\varepsilon$，

$$\lim_{t\to -\varepsilon^-} h_\varepsilon^k(t)=0\ ,\ \ \lim_{t\to -\varepsilon^+} h_\varepsilon^k(t)=0$$

Then: $\lim_{t\to -\varepsilon^-} h_\varepsilon^k(t)=\lim_{t\to -\varepsilon^+} h_\varepsilon^k(t)$.

To sum up, $h_\varepsilon^k(t)$ is continuous in $t=0$， $t=-\varepsilon$, $h_\varepsilon^k(t)$ is continuous in R .

(2) To prove $\forall t\in R, h_\varepsilon^m(t)\ge h^m(t)$ ， Just prove that $\Delta h(t):=h_\varepsilon^k(t)-h^k(t)\ge 0$ ， By definition:

$$\Delta h(t):=h_\varepsilon^k(t)-h^k(t):=\begin{cases}\Delta h_1(t)=0, & t\le -\varepsilon\\ \Delta h_2(t)=\varepsilon^{k(1-b)}b^{-k}(t+\varepsilon)^{kb}, & -\varepsilon<t\le 0\\ \Delta h_3(t)=(t+\varepsilon b^{-1})^k-t^k, & t>0\end{cases}$$

(i) If $t\le -\varepsilon$, $\Delta h(t)=\Delta h_1(t)=0\ge 0$ is hold;

(ii) If $-\varepsilon<t\le 0$， then $\varepsilon>0$， $b>0$，且$t+\varepsilon>0$，于是:

$$\Delta h(t)=\Delta h_2(t)=\varepsilon^{k(1-b)}b^{-k}(t+\varepsilon)^{kb}>0$$

(iii) If $t>0$, because $\varepsilon>0$， $b>0$， then $\varepsilon b^{-1}>0$:

$$\Delta h(t)=\Delta h_3(t)=(t+\varepsilon b^{-1})^k-t^k>0$$

To sum up, $\Delta h(t):=h_\varepsilon^k(t)-h^k(t)\ge 0$, then $\forall t\in R, h_\varepsilon^m(t)\ge h^m(t)$

By the construction of $h_\varepsilon^k(t)$,

(i) If $t\le -\varepsilon$, $\lim_{\varepsilon\to 0} h_\varepsilon^k(t)=0$ is hold;

(ii) If $-\varepsilon<t\le 0$, then，

$$\frac{dh_\varepsilon^k(t)}{dt}=k\varepsilon^{k(1-b)}b^{1-k}(t+\varepsilon)^{kb-1}\ ,$$

Because $\varepsilon>0$, $b>0$, and $t+\varepsilon>0$,then：

$$\frac{dh_\varepsilon^k(t)}{dt}=k\varepsilon^{k(1-b)}b^{1-k}(t+\varepsilon)^{kb-1}>0$$

Namely smooth function $h_\varepsilon^k(t)$ is constant monotonous increase in $-\varepsilon<t\le 0$, then

$$0=h(-\varepsilon)<h_\varepsilon^k(t)\le h(0)=\varepsilon^k b^{-k}$$

then

$$0\le \lim_{\varepsilon\to 0} h_\varepsilon^k(t)\le \lim_{\varepsilon\to 0}\varepsilon^k b^{-k}=0$$

$\lim_{\varepsilon\to 0} h_\varepsilon^k(t)=0$ is hold.

(iii) If $t>0$:

$$\lim_{\varepsilon\to 0} h_\varepsilon^k(t)=\lim_{\varepsilon\to 0}(t+\varepsilon b^{-1})^k=t^k=h^k(t)$$

In summary, $\forall t \in R, \lim_{\varepsilon \to 0} h_{\varepsilon}^{m}(t) = h^{m}(t)$

(3)By the construction of $h_{\varepsilon}^{k}(t)$，notes：

$$h_{\varepsilon}^{k}(t) = \begin{cases} h_1(\varepsilon,t) = 0 & t < -\varepsilon \\ h_2(\varepsilon,t) = \varepsilon^{k(1-b)} b^{-k} (t+\varepsilon)^{kb} & -\varepsilon \le t < 0 \\ h_3(\varepsilon,t) = (t+\varepsilon b^{-1})^{k} & t \ge 0 \end{cases}$$

First of all:

$$h_1(t,\varepsilon) = 0 < h_2(t,\varepsilon) \le \varepsilon^{k} b^{-k} \le h_3(t,\varepsilon)$$

Followed by:

$$\begin{aligned} \frac{dh_2(t,\varepsilon)}{d\varepsilon} &= k\varepsilon^{k(1-b)} b^{1-k} (t+\varepsilon)^{kb-1} + k(1-b)\varepsilon^{k(1-b)-1} b^{-k} (t+\varepsilon)^{kb} \\ &= k\varepsilon^{k(1-b)-1} b^{-k} (t+\varepsilon)^{kb-1} (\varepsilon b + (1-b)(t+\varepsilon)) \\ &= k\varepsilon^{k(1-b)-1} b^{-k} (t+\varepsilon)^{kb-1} (t+\varepsilon - bt)) \end{aligned}$$

Due to $\varepsilon > 0$，$b > 0$，$-\varepsilon < t \le 0$，then：

$$\frac{dh_2(t,\varepsilon)}{d\varepsilon} > 0$$

and

$$\frac{dh_3(t,\varepsilon)}{d\varepsilon} = b^{-1}(t+\varepsilon b^{-1})^{k-1} > 0$$

In summary, smooth function $h_{\varepsilon}^{k}(t)$ is Constant monotonous increase in $\varepsilon \ge 0$ which means $0 < \varepsilon_1 < \varepsilon_2$,then $\forall t \in R, h_{\varepsilon_1}^{k}(t) \le h_{\varepsilon_2}^{k}(t)$ 。

Theorem 2.2.2 if $\frac{1}{b} < k < +\infty$，$\varepsilon > 0$，$h_{\varepsilon}^{k}(t)$ is continuous differentiable，and

$$\frac{dh_{\varepsilon}^{k}(t)}{dt} = \begin{cases} 0 & t \le -\varepsilon \\ k\varepsilon^{k(1-b)} b^{1-k} (t+\varepsilon)^{kb-1} & -\varepsilon < t \le 0 \\ k(t+\varepsilon b^{-1})^{k-1} & t > 0 \end{cases}$$

Consider the low-order penalty function

$$\widehat{P}_{\varepsilon}^{k}(x,\pi) = \sum_{i=1}^{m} f_i(x) + \pi \sum_{i=1}^{l} h_{\varepsilon}^{k}(g_i(x)) + \pi \sum_{i=1}^{m} h_{\varepsilon}^{k}(f_i(x) - f_i(\hat{x}))$$

Theorem 2.2.3 For any $\frac{1}{b} < k < 1$，$\varepsilon > 0$，$b > 0$，and $f_i(x), i \in M$，$g_i(x), i \in I$ is continuous，the following properties hold:

(1) $\widehat{P}_{\varepsilon}^{k}(x,\pi)$ is continuous in R^{n} ;

(2) $\forall x \in R^{n}, \widehat{P}_{\varepsilon}^{k}(x,\pi) \ge \widehat{P}^{k}(x,\pi)$;

(3) $\forall x \in R^{n}, \lim_{\varepsilon \to 0} \widehat{P}_{\varepsilon}^{k}(x,\pi) = \widehat{P}^{k}(x,\pi)$;

(4)if $0 < \varepsilon_1 \le \varepsilon_2$, then $\forall x \in R^{n}, \widehat{P}_{\varepsilon_1}^{k}(x,\pi) \le \widehat{P}_{\varepsilon_2}^{k}(x,\pi)$, $\forall x \in R^{n}, \hat{P}_{\varepsilon_1}^{k}(x,\pi) \le \hat{P}_{\varepsilon_2}^{k}(x,\pi)$.

Theorem 2.2.4 If $\frac{1}{b} < k < 1$，$b > 0$，$\varepsilon > 0$，and $f_i(x), i \in M$，$g_i(x), i \in I$ is continuous differentiable，then $\widehat{P}_\varepsilon^k(x,\pi)$ is continuous differentiable in R^n.

$$\nabla_x \widehat{P}_\varepsilon^k(x,\pi) = \sum_{i=1}^{m} \nabla_x f_i(x) + \pi \sum_{i=1}^{l} \nabla_x h_\varepsilon^k(g_i(x)) + \pi \sum_{i=1}^{m} \nabla_x h_\varepsilon^k(f_i(x) - f_i(\hat{x}))$$

3 A MOSQP algorithm based on the low-order smooth penalty function

3.1 Quadratic programming sub-problem construction

For the initial value diffusion phase, we construct the following problems for each objective function of the multi-objective optimization problem:

$$\begin{aligned} &(\mathrm{P}_i) \quad \min\ f_i(x), i \in M = \{1, \cdots, m\} \\ &s.t. \quad g_i(x) \le 0, i \in I = \{1, 2, \cdots a\} \end{aligned}$$

The value function we used at this stage is as: $f_i(x)$.

The search direction is obtained by solving the following quadratic planning subproblems:

$$\begin{aligned} &\mathbf{QP}_i(x^k, B^k) \quad \min \frac{1}{2}(d^j)^T B^j d^j + \nabla f_i(x^j)^T d^j, i \in M = \{1, \cdots, m\} \\ &s.t.\ \nabla g_i(x^j)^T d + g_i(x^j) \le 0 \quad i \in I \end{aligned}$$

For the multi-objective optimization problem Pareto optimal solution stage, we consider all the objective functions and take each iterative initial value point as the reference point $\hat{x}$，Constructing the problem(P2)：

$$\begin{aligned} &(\mathrm{P2}) \quad \min \sum_{i=1}^{m} f_i(x) \\ &s.t. \quad f_i(x) \le f_i(\hat{x}), i \in M = \{1, \cdots, m\} \\ &\qquad\quad g_i(x) \le 0, i \in I = \{1, 2, \cdots a\} \end{aligned}$$

The corresponding low-order penalty function is:

$$\widehat{P}_\varepsilon^k(x,\pi) = \sum_{i=1}^{m} f_i(x) + \pi \sum_{i=1}^{l} h_\varepsilon^k(g_i(x)) + \pi \sum_{i=1}^{m} h_\varepsilon^k(f_i(x) - f_i(\hat{x}))$$

The search direction is obtained by solving the following quadratic planning subproblems：