Lecture 4. Tikhonov method for equations with convex functional. Local strict convexity of the Tikhonov functional for non-linear operators.

January 14, 2010

1 Tikhonov method for equations with convex functional

Tikhonov scheme allows to construct approximation methods for solution of equation $F(x) = 0$ not only for special case of operator F, when $F = Ax - f$.

In this lesson we will analyze more common case when functional Φ is convex, but not necessary quadratic. As before, we assume that $F = F(N_1, N_2)$. If equation have form $Ax = f$, then corresponding functional

$$\Phi(x) = \frac{1}{2}||Ax - f||_{H_2}^2$$

is convex and class of operator equations which we are interested in, is not empty. This class is extension of class $F(N_1, 0)$. Differentiating Tikhonov functional

$$\Phi_\alpha(x) = \frac{1}{2}||Ax - f||_{H_2}^2 + \frac{1}{2}\alpha||x - x_0||_{H_1}^2$$

gives as

$$\Phi'_\alpha(x) = \Phi'(x) + \alpha(x - x_0), x \in H_1.$$

Because functional Φ is convex, we have

$$(\Phi'(x) - \Phi'(y), x - y)_{H_1} \geq 0, \forall x, y \in H_1.$$

Thus,

$$(\Phi'_\alpha(x) - \Phi'_\alpha(y), x - y)_{H_1} \geq \alpha||x - y||_{H_1}^2, \forall x, y \in H_1. \tag{1}$$

Equation (2) says that functional Φ_α is strongly convex in H_1 for $\forall\alpha > 0$. from theory of convex optimization is known that in this case exists unique global minimizer of this problem which realizes minimum of functional Φ_α in H_1.

$$\Phi_\alpha(x_\alpha) = \inf_{x \in H_1} \Phi_\alpha(x), \alpha > 0. \tag{2}$$

In optimization theory there exists a lot of minimization methods for strongly convex functionals. In all algorithms we face out with two difficulties:

- first problem is that there are not convergence estimates of x_α to x^* when $\alpha \to 0$.

- second problem is that iterative processes are less effective with reducing of parameter α since these processes requires increasing of optimization iterations when $\alpha \to 0$.

Let us clarify this on one example when we apply gradient method for minimizing of functional Φ_α:

$$x_0 \in H_1 : x_{n+1} = x_n - \gamma \Phi'_\alpha(x_n), n = 0, 1, ... \tag{3}$$

Because of (3) and necessary minimum condition

$$\Phi'_\alpha(x_\alpha) = \Phi'(x_\alpha) + \alpha(x_\alpha - x_0) = 0, n = 0, 1, ...$$

we have

$$\begin{aligned} ||x_{n+1} - x_\alpha||^2_{H_1} &= ||x_n - x_\alpha - \gamma(\Phi'_\alpha(x_n) - \Phi'(\alpha(x_\alpha)))||^2_{H_1} \\ &= ||x_n - x_\alpha||_{H_1} - 2\gamma(\Phi'_\alpha(x_n) - \Phi'_\alpha(x_\alpha), x_n - x_\alpha)_{H_1} \\ &+ \gamma^2||\Phi'_\alpha(x_n) - \Phi'_\alpha(x_\alpha)||^2_{H_1}. \end{aligned} \tag{4}$$

Remind that

$$\Phi'(x) = F^{*'}(x)F(x), \quad x \in H_1.$$

Following estimate takes place

$$\begin{aligned} ||\Phi'_\alpha(x_n) - \Phi'_\alpha(x_\alpha)||^2_{H_1} &\le \\ &\le ||\Phi'(x_n) - \Phi'(x_\alpha)||_{H_1} + \alpha||x_n - x_\alpha||_{H_1} \\ &\le ||F^{'*}(x_n)||_{L(H_1,H_2)}||F(x_n) - F(x_\alpha)||_{H_2} \\ &+ ||F^{'*}(x_n) - F^{'*}(x_\alpha)||_{L(H_2,H_1)}||F(x_\alpha)||_{H_2} \\ &+ \alpha||x_n - x_\alpha||_{H_1} \\ &\le (N_1^2 + N_2||F(x_\alpha)||_{H_2} + \alpha)||x_n - x_\alpha||_{H_1}. \end{aligned} \tag{5}$$

By definition of x_α

$$\Phi_\alpha(x_\alpha) = \frac{1}{2}||F(x_\alpha)||^2_{H_2} + \frac{\alpha}{2}||x_\alpha - x_0||^2_{H_1} \le \Phi_\alpha(x^*_{x_o} - x_0)^2_{H_1}$$

and thus

$$||F(x_\alpha)||_{H_2} \le \sqrt{\alpha}\Phi_\alpha(x^*_{x_o} - x_0)_{H_1} \tag{6}$$

From estimates (5) and (6) follows that

$$\begin{aligned} ||\Phi'_\alpha(x_n) - \Phi'_\alpha(x_\alpha)||^2_{H_1} &\le L||x_n - x_\alpha||_{H_1}, \\ L &= N_1^2 + N_2||x^*_{x_0} - x_0||_{H_1}\sqrt{\alpha} + \alpha. \end{aligned} \tag{7}$$

Substituting estimates (5) and (7) in (4) and using (2) we can get

$$||x_{n+1} - x_\alpha||^2_{H_1} \le \sqrt{1 - 2\gamma\alpha + \gamma^2 L^2}||x_n - x_\alpha||_{H_1}. \tag{8}$$

We see that for $\forall \gamma > 0$ in (8) $1 - 2\gamma\alpha + \gamma^2 L^2 > 0$. Value of γ we should choose like that expression $1 - 2\gamma\alpha + \gamma^2 L^2$ should be small. We can require that

$$1 - 2\gamma\alpha + \gamma^2 L^2 < 1.$$

For this it is sufficient that following condition should be fullfilled

$$0 < \gamma < \frac{2\alpha}{L^2}. \tag{9}$$

In the case that (9) is true we have

$$\begin{gathered} ||x_n - x_\alpha||^2_{H_1} \leq \sqrt{1 - 2\gamma\alpha + \gamma^2 L^2}||x_0 - x_\alpha||_{H_1} q^n(\alpha), \\ q(\alpha) = \sqrt{1 - 2\gamma\alpha + \gamma^2 L^2}, \quad q(\alpha) \in (0, 1). \end{gathered} \tag{10}$$

From (10) follows that defined by (3) sequence x_n converges to x_α as geometric progression. Then minimal value its achieved at $\gamma = \alpha L^{-2}$ and corresponding minimal value of $1 - 2\gamma\alpha + \gamma^2 L^2$ is equal to $1 - \alpha^2/L^2$. Thus, when $\alpha \to 0$ then $q(\alpha) \to 1$ for any choice of γ from condition (9). Because of this we need to make more iterations in minimizing of our functional.

1.1 Convexity of Tikhonov functional for linear operators

Let H_1 and H_2 be two real valued Hilbert spaces. Denote norms in these spaces as $\|\bullet\|_1$ and $\|\bullet\|_2$. Let the operator $F : H_1 \to H_2$. Consider a general equation

$$F(x) = 0, x \in H_1. \tag{11}$$

An important case of this equation is

$$F(x) = Ax - f, A \in \mathbb{L}(H_1, H_2); x \in H_1, f \in H_2. \tag{12}$$

Thus, we will discuss now the linear operator equation

$$Ax = f, x \in H_1, A \in \mathbb{L}(H_1, H_2), f \in H_2. \tag{13}$$

In other words, $A : H_1 \to H_2$ is a bounded linear operator. In the case of ill-posed problems A is a compact operator. Therefore, a bounded inverse A^{-1} does not exist. Thus, we need to consider the Tikhonov functional Φ_α to approximately solve equation (3), where

$$\Phi_\alpha(x) = \frac{1}{2}\|Ax - f\|_2^2 + \frac{\alpha}{2}\|x - x_0\|_1^2. \tag{A}$$

Here x_0 is a certain a priori chosen vector. In computational practice usually x_0 is a good approximation for the correct solution.

We now establish the strict convexity of $\Phi_\alpha(x)$. This is an important property. Indeed, it is well known that a strictly convex functional can have at most one point of minimum. Now, if $\dim H_1 < \infty$, then this minimum exists. Indeed, it is clear that

$$\lim_{\|x\|_1 \to \infty} \Phi_\alpha(x) = \infty.$$

On the other hand,

$$\Phi_\alpha(0) = \frac{\|f\|_2^2}{2} + \frac{\alpha \|x_0\|_1^2}{2}.$$

In other words, $\Phi_\alpha(0)$ has a certain bounded value. Hence, by the Weierstrass theorem a minimizer x_α of $\Phi_\alpha(x)$ exists and by Theorem 1 this minimizer is unique: the uniqueness is because $\Phi_\alpha(x)$ is strictly convex.

Definition 1. Let $U \subseteq H_1$ be convex set and the operator $G : U \to H_2$ (G is not necessary linear). The operator G has the Frechet derivative $G'(x) \in L(H_1, H_2)$ at the point $x \in U$ if the following representation is valid

$$G(x+h) - G(x) = G'(x)(h) + o_x(\|h\|_1), \forall h \in H_1, \tag{100}$$

where

$$\lim_{\|h\|_1 \to 0} \frac{\|o_x(\|h\|_1)\|_2}{\|h\|_1} = 0. \tag{101}$$

In other words, the difference $G(x+h) - G(x)$ is a linear operator $G'(x)(h)$ plus a function, whose norm us much smaller than $\|h\|_1$ for sufficiently small $\|h\|_1$. This is the full similarity with the conventional definition of the differential.

Furthermore, it follows from (100) and (101) that in order to figure out the Frechet derivative $G'(x)(h)$ at the point x, one should consider the difference $G(x+h) - G(x)$ and consider the linear, with respect to h, part of this difference. So, this part is a good candidate for being the Frechet derivative: do not forget that we also need to prove (101).

Definition. Let $U \subseteq H_1$ be convex set and let the functional $\Phi : U \to \mathbb{R}$ has the Frechet derivative $\Phi'(x)$ for all $x \in U$. This functional is called strictly convex on the set U with the strict convexity parameter κ if

$$\Phi(x+y) - \Phi(x) - \Phi'(x)(y) \geq \kappa \|y\|_1^2, \forall x, y \in U \text{ such that } x + y \in U. \tag{B}$$

Theorem 1. *Suppose that the operator A in (13) is linear and bounded. Then the functional Φ_α is strictly convex on the space H_1 with the strict convexity parameter $\alpha/2$.*

Proof. First, we consider the Frechet derivative $\Phi'_\alpha(x)$ of the functional $\Phi_\alpha(x)$. So, by the above rule, we need to consider the difference $\Phi_\alpha(x+h) - \Phi_\alpha(x)$, single out its linear part (with respect to h) and prove (101) for the rest. Let $(,)_1$ and $(,)_2$ be scalar products in H_1 and H_2 respectively. Since we have real valued Hilbert spaces, then $(a,b)_k = (b,a)_k, \forall a, b \in H_k, k = 1, 2$. Hence,

$$\Phi_\alpha(x+h) - \Phi_\alpha(x) = \frac{1}{2}(Ax + Ah, Ax + Ah)_2 + \frac{\alpha}{2}(x+h, x+h)_1$$

$$-\frac{1}{2}(Ax,Ax)_2-\frac{\alpha}{2}(x,x)_1 \tag{C}$$

$$=\left[(Ax,Ah)_2+\alpha(x,h)_1\right]+\left[\frac{1}{2}(Ah,Ah)_2+\frac{\alpha}{2}(h,h)_1\right].$$

Since A is bounded operator, than

$$\left|\frac{1}{2}(Ah,Ah)_2+\frac{\alpha}{2}(h,h)_1\right|\leq\frac{1}{2}(\|A\|+\alpha)\|h\|_1^2.$$

Hence,

$$\begin{aligned}&\lim_{\|h\|_1\to 0}\frac{1}{\|h\|_1}\left\{\left[\frac{1}{2}(Ah,Ah)_2+\frac{\alpha}{2}(h,h)_1\right]\right\}\\ \leq\ &\frac{1}{2}\lim_{\|h\|_1\to 0}(\|A\|+\alpha)\|h\|_1=0.\end{aligned}$$

Therefore, it follows from (100), (101) and (C) that

$$\begin{aligned}\Phi'_\alpha(x)(h)&=(Ax,Ah)_2+\alpha(x,h)_1\\ \text{or }\Phi'_\alpha(x)(h)&=((AA^*+\alpha)x,h)_1.\end{aligned}$$

Hence, $\Phi'_\alpha(x)=(AA^*+\alpha I)x$, where I is the identity operator on H_1.

Now, we are ready to use (B) to prove the strict convexity of Φ_α. We have

$$\begin{aligned}&\Phi_\alpha(x+y)-\Phi_\alpha(x)-\Phi'_\alpha(x)(y)\\ &=\frac{1}{2}(Ax+Ay,Ax+Ay)_2+\frac{\alpha}{2}(x+y,x+y)_1\\ &-\frac{1}{2}(Ax,Ax)_2-\frac{\alpha}{2}(x,x)_1-(Ax,Ay)_2-\alpha(x,y)_1\\ &=\frac{1}{2}(Ay,Ay)_2+\frac{\alpha}{2}(y,y)_1\geq\frac{\alpha}{2}\|y\|_1^2.\end{aligned}$$

We have used here the fact that $(Ay,Ay)_2\geq 0$. $\square$

2 Existence of minimizer

As before we assume that there exists Frechet derivative of functional $\Phi_\alpha(x)$ and $F\in(N_1,N_2)$.

We assume that

$$||F(x^*)||\leq\delta, 0<\delta<1. \tag{14}$$

We also assume that the operator F has the Fechet derivative $F'(x)$ for $x\in V_1(x^*)=\{||x-x^*||<1\}$ and this derivative is Lipschitz continuous or

$$||F'(x)||\leq N_1, ||F'(x)-F'(y)||<N_2||x-y||. \tag{15}$$

We now assume that

$$\begin{aligned} \|x_{glob} - x^*\| &\leq \delta^{\mu_1}, \mu_1 = const. \in (0,1), &(16)\\ \alpha &= \delta^{\mu_2}, \mu_2 = const. \in (0, \min(\mu_1, 2(1-\mu_1))) &(17)\end{aligned}$$

Lemma *A minimizer x_α of the functional $\Phi_\alpha(x)$ on the space H_1 exists for any value of the regularization parameter α. For any $r > 0$ denote $V_r(x_\alpha) = \{x \in H_1 : \|x - x_\alpha\| < r\}$. Assume that conditions (16), (17) hold. Then $x_{glob} \in V_{\sqrt{2}\delta^{\mu_1}}(x_\alpha)$ and $x^* \in V_{(1+\sqrt{2})\delta^{\mu_1}}(x_\alpha)$. Let $\beta_1 \in (0,1)$ be any number. Then there exists a sufficiently small number $\delta_0 = \delta_0(\mu_1, \mu_2, \beta_1) \in (0,1)$ such that if $\delta \in (0, \delta_0)$, then $x^*, x_{glob} \in V_{\beta_1 \alpha}(x_\alpha)$.*

Proof.

Since $\dim H < \infty$, then $\lim_{\|x\| \to \infty} \Phi_\alpha(x) = \infty$ implies the existence of a minimizer x_α.

Since $\Phi_\alpha(x_\alpha) \leq \Phi_\alpha(x^*)$ then

$$\Phi_\alpha(x^*) = \Phi(x^*) + \frac{\alpha}{2}||x^* - x_{glob}|| = \frac{1}{2}||F(x^*)||^2 + \frac{\alpha}{2}||x^* - x_{glob}|| \leq \frac{\delta^2 + \alpha\delta^{2\mu_1}}{2}, \qquad (18)$$

and then

$$\Phi_\alpha(x_\alpha) \leq \Phi_\alpha(x^*) = \frac{\delta^2 + \alpha\delta^{2\mu_1}}{2} \qquad (19)$$

and thus

$$\Phi_\alpha(x_\alpha) \leq \frac{\delta^2 + \alpha\delta^{2\mu_1}}{2} < \alpha\delta^{2\mu_1} \qquad (20)$$

Why $\frac{\delta^2+\alpha\delta^{2\mu_1}}{2} < \alpha\delta^{2\mu_1}$?

Since $\alpha = \delta^{\mu_2}$, then

$$\frac{\alpha\delta^{2\mu_1}}{\delta^2} = \delta^{\mu_2 - 2(1-\mu_1)}.$$

By (17) $\mu_2 < 2(1-\mu_1)$. Hence $\mu_2 - 2(1-\mu_1) < 0$. Hence,

$$\lim_{\delta \to 0} \delta^{\mu_2 - 2(1-\mu_1)} = \infty. \qquad (D)$$

But in Lemma we talk about sufficiently small δ_0. Hence, it follows from (D) that if δ_0 is sufficiently small, then for $\delta \in (0, \delta_0)$ we have

$$\frac{\alpha\delta^{2\mu_1}}{\delta^2} = \delta^{\mu_2 - 2(1-\mu_1)} > 1 \to \alpha\delta^{2\mu_1} > \delta^2.$$

Then, by the definition of $\Phi_\alpha(x_\alpha)$ we have

$$\Phi_\alpha(x_\alpha) = \Phi(x_\alpha) + \frac{\alpha}{2}||x_\alpha - x_{glob}||^2 = \frac{1}{2}||F(x_\alpha)||^2 + \frac{\alpha}{2}||x_\alpha - x_{glob}||^2 < \alpha\delta^{2\mu_1}. \qquad (21)$$

It follows from this equation that

$$\frac{1}{2}||F(x_\alpha)||^2 < \alpha\delta^{2\mu_1}. \tag{22}$$

and

$$\frac{\alpha}{2}||x_\alpha - x_{glob}||^2 < \alpha\delta^{2\mu_1}. \tag{23}$$

From the last inequality we get that

$$||x_\alpha - x_{glob}|| < \sqrt{2}\delta^{\mu_1}. \tag{24}$$

Using this estimate, we obtain

$$\begin{aligned} ||x_\alpha - x_{glob}|| &\leq ||x_\alpha - x_{glob}|| + ||x_0 - x^*|| \\ &< \sqrt{2}\delta^{\mu_1} + \delta^{\mu_1} \\ &= (1+\sqrt{2})\delta^{\mu_1} < \beta_1\alpha = \beta_1\delta^{\mu_2}. \end{aligned} \tag{25}$$

Why $(1+\sqrt{2})\delta^{\mu_1} < \beta_1\alpha = \beta_1\delta^{\mu_2}$**?**

Since $\mu_2 < \mu_1$, then

$$\lim_{\delta\to 0}\frac{\beta_1\delta^{\mu_2}}{(1+\sqrt{2})\delta^{\mu_1}} = \infty$$

and δ is sufficiently small. Thus

$$\frac{\beta_1\delta^{\mu_2}}{(1+\sqrt{2})\delta^{\mu_1}} > 1$$

and thus

$$\beta_1\delta^{\mu_2} > (1+\sqrt{2})\delta^{\mu_1}.$$

Theorem *Assume that conditions (16), (17) hold. Then there exists numbers* $\beta_1 = \beta_1(N_1, N_2) \in (0,1)$ *and* $\delta_1 = \delta_1(\mu_1, \mu_2, N_2, \beta_1) \in (0,1)$ *depending only on listed parameters such that if* $\rho = \beta_1\alpha$*, then for any* $\delta \in (0, \delta_1)$ *the functional* J_α *is strictly convex in the neighborhood* $V_\rho(x_\alpha)$ *of the point* x_α *with the strict convexity parameter* $\kappa = \alpha/4$*. Furthermore, by Lemma 2.1 points* $x_{glob}, x^* \in V_\rho(x_\alpha)$*.*

Proof. Let $\beta_1 \in (0,1)$ be the number which we will choose below in this proof, $\rho = \beta_1\alpha$ and $x, y \in V_\rho(x_\alpha)$ be two arbitrary points. By (2.9)

$$\begin{aligned} \left(J_\alpha'(x) - J_\alpha'(y), x-y\right) &= \alpha\|x-y\|^2 + \left(F'^*(x)F(x) - F'^*(y)F(y), x-y\right) \\ &= \alpha\|x-y\|^2 + \left(F'^*(x)F(x) - F'^*(x)F(y), x-y\right) \\ &\quad + \left(F'^*(x)F(y) - F'^*(y)F(y), x-y\right). \end{aligned} \tag{26}$$

Denote $A_1 = \left(F'^*(x)F(x) - F'^*(x)F(y), x-y\right)$, $A_2 = \left(F'^*(x)F(y) - F'^*(y)F(y), x-y\right)$ and estimate A_1, A_2 from the below.

Since $A_1 = A_1 - \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right) + \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right)$, then we have

$$
\begin{aligned}
A_1 &= \left(F'^*\left(x\right)F\left(x\right) - F'^*\left(x\right)F\left(y\right), x-y\right) - \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right) \\
&+ \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right) \\
&= \left(F'^*\left(x\right)\left(F\left(x\right) - F\left(y\right), x-y\right) - \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right) \\
&+ (F'^*(x)F'(x)(x-y), x-y),
\end{aligned}
$$

then using main theorem of calculus

$$F(x) - F(y) = \int\limits_0^1 \left(F'\left(y + \theta\left(x-y\right)\right)\right)\left(x-y\right)d\theta$$

we can rewrite A_1 as

$$
\begin{aligned}
A_1 &= \left(F'^*\left(x\right)\int\limits_0^1 \left(F'\left(y+\theta\left(x-y\right)\right) - F'\left(x\right)\right)\left(x-y\right)d\theta, x-y\right) \\
&\quad + \left(F'^*\left(x\right)F'\left(x\right)\left(x-y\right), x-y\right).
\end{aligned}
$$

We obtain

$$
\begin{aligned}
&|F'^*(x)\int\limits_0^1 [F'(y+\theta(x-y)) - F'(x)](x-y)d\theta, x-y| \\
\leq\ & \left\|F'\left(x\right)\right\| \int\limits_0^1 \left\|\left[F'\left(y+\theta\left(x-y\right)\right) - F'\left(x\right)\right]\left(x-y\right)\right\| d\theta \cdot \|x-y\| \\
\leq\ & N_1 N_2 \int\limits_0^1 (y+\theta(x-y)-x)(x-y)d\theta \cdot \|x-y\| \\
\leq\ & \frac{1}{2} N_1 N_2 \left\|x-y\right\|^3.
\end{aligned}
$$

$$z = \int\limits_0^1 [F'(y+\theta(x-y)) - F'(x)](x-y)d\theta.$$

Hence,

$$\left|\left(F'^*(x)\int\limits_0^1 [F'(y+\theta(x-y)) - F'(x)](x-y)d\theta, x-y\right)\right|$$

$$= |(F'^{*}(x)z, x-y)| \le \|F'^{*}(x)\| \cdot \|z\| \cdot \|x-y\|$$
$$= \|F'(x)\| \cdot \|z\| \cdot \|x-y\| .$$

We have used here that for every bounded linear operator A we have $\|A^{*}\| = \|A\|$.

Also,

$$(F'^{*}(x) F'(x)(x-y), x-y) = (F'(x)(x-y), F'(x)(x-y))_2 = \|F'(x)(x-y)\|_2^2 \ge 0.$$

Hence, $A_1 \ge N_1 N_2 \|x-y\|^3 / 2$. Now we estimate A_2,

$$|A_2| \le \|F(y)\|_2 \|F'(x) - F'(y)\| \|x-y\| \le N_2 \|x-y\|^2 \|F(y)\|_2 .$$

Since

$$\|F(y)\|_2 \le \|F(y) - F(x_\alpha)\|_2 + \|F(x_\alpha)\|_2 \le N_1 \|y - x_\alpha\| + \|F(x_\alpha)\|_2 , \tag{27}$$

since

$$||F(y) - F(x_\alpha)||_2 \le (\max_{x \in V_1(x^*)} ||F'(x)||)||y - x_\alpha|| \le N_1||y - x_\alpha||, \tag{28}$$

then

$$|A_2| \le N_2 \|x-y\|^2 (N_1 \|y - x_\alpha\| + \|F(x_\alpha)\|_2) . \tag{29}$$

By (14), (15) and Lemma $\|F(x_\alpha)\|_2 \le \|F(x_\alpha) - F(x^*)\|_2 + \delta \le \alpha\beta_1 N_1 + \delta$.

Hence,

$$A_2 \ge -N_2 \|x-y\|^2 (N_1 \|y - x_\alpha\| + \alpha\beta_1 N_1 + \delta). \tag{30}$$

Combining this with (26) and the above estimate for A_1, we obtain

$$(\Phi'_\alpha(x) - \Phi'_\alpha(y), x-y) \ge \tag{31}$$
$$\|x-y\|^2 \left[\alpha - \frac{N_1 N_2}{2} \|x-y\| - N_1 N_2 \|y - x_\alpha\| - N_2 (N_1 \alpha \beta_1 + \delta) \right] .$$

We have by Lemma

$$\frac{||x-y||}{2} = \frac{||x - x_\alpha + x_\alpha - y||}{2} \le \frac{\alpha\beta_1 + \alpha\beta_1}{2} = \alpha\beta_1. \tag{32}$$

and thus

$$\begin{aligned} & N_1 N_2 \frac{\|x-y\|}{2} + N_1 N_2 ||y - x_\alpha|| + N_2 (N_1 \alpha\beta_1 + \delta) \\ & \le \alpha\beta_1 N_2 N_1 + \alpha\beta_1 N_2 N_1 + \alpha\beta_1 N_2 N_1 + N_2 \delta \\ & \le 3\alpha\beta_1 N_2 N_1 + N_2 \delta. \end{aligned} \tag{33}$$

Choose $\beta_1 = \beta_1(N_1, N_2) \in (0,1)$ such that $3\beta_1 N_2 N_1 \le 1/4$. Given this β_1, choose $\delta_1 = \delta_1(\mu_1, \mu_2, N_1, N_2) \in (0,1)$ so small that

$$N_2 \delta < \delta^{\mu_2}/4 = \alpha/4 \tag{34}$$

and $2\delta^{\mu_1} < \beta_1 \delta^{\mu_2} = \beta_1 \alpha, \forall \delta \in (0, \delta_1)$. Then we have that $\alpha/4 + \alpha/4 = \alpha/2$ in (33) and thus from (31) we have $\alpha - \alpha/2 = \alpha/2$.

□