Dear Professor Li:

I have very carefully examined the convergence proof for both the first order and second order necessary conditions. So far, I feel confident that the results in regard to 2nd order conditions are correct and can be presented in the thesis (without the proof). I have also revised the thesis upon your latest comments and hopefully only few corrections will be needed.

I have several questions about the overall proof and I hope you can give me some suggestions after reviewing my questions. Thank you.

(1) As I brought to your attention in the last email, in our context, gradient $g(x)$ is no longer uniformly continuous thus we can not say when $\|x_{m_i} - x_{l_i}\| \to 0$, this implies $\|g_{m_i} - g_{l_i}\| \le \epsilon_2$ for sufficiently large i. If we think of strict complementarity condition as implying, in a limit, it is not possible for x_{m_i} and x_{l_i} to be separated by non-differentiable hyperplanes when $\|x_{m_i} - x_{l_i}\| \to 0$, i.e., they must reside in the same orthant, then we can claim $\|g_{m_i} - g_{l_i}\| \le \epsilon_2$ since there is no sign changes. Or we may need other assumptions to make this true?

(2) In your original paper, I am confused about your definition of $\phi_k^*[d_k]$. In page 6, equation 2.12, $\phi_k^*[d_k]$ is used to denote the minimum value of $\phi_k(s)$ along the direction d_k within the feasible trust region, i.e., (I use scaling in our context to avoid ambiguity)

$$\phi_k^*[d_k] := \phi(\tau^* d_k) := \min\left\{\phi_k(\tau d_k) : \left\|\tau D_k^{-\frac{1}{2}} d_k\right\| \le \Delta_k, x_k + \tau d_k \in diff(\mathcal{F})\right\}$$

while in page 19 in proving $\liminf_{k\to\infty} \frac{\phi_k(\beta_k^*[p_k])}{\phi_k^*[p_k]} \ge 1$, you use $\phi_k^*[p_k] = \tau_k^* g_k^T p_k + \frac{1}{2}(\tau^*)^2 p_k^T M_k p_k$ where $\tau^* = \min\left\{1, \beta_k^1\right\}$. My question is this τ^* is the minimizer for the $\phi(\tau)$ function defined in Lemma 3.7, which forces $\tau \in \left[0, \min\left\{1, \beta_k^1\right\}\right]$, it is NOT the minimizer for the trust region subproblem? For $\phi_k^*[p_k]$, is it as simple as $\phi_k^*(p_k)$ with $\tau^* = 1$? As you can see in the attached file with full proof (file name:ThesisTemplateFinalSept3.pdf), in Lemma 4.7, I use $\Phi(\tau) := \phi_k(\tau p_k), \tau \in \left[0, \min\left\{1, \beta_k^1\right\}\right]$ to make the distinction from $\phi_k^*[d_k]$. Though it has been proved $\liminf_{k\to\infty} \beta_k^1 \ge 1$ in Lemma 4.9, I don't know if my observation is correct?

(3) Throughout the analysis, now I use $\psi(d_k)$ to indicate the objective function for the trust region subproblem. $\phi(d_k)$ is used to measure the decrease in the objective values. Assumption 4 and 6 are used to connect the decrease by taking step d_k derived from trust region subproblem to $\phi_k^*[p_k]$ and $\phi_k^*[-D_k g_k]$. I think in the draft paper you gave me on Stable Local Volatility Function Calibration, in Page 8, you sufficient decrease assumption

$$\phi_k(d_k) \le \beta_g \phi_k^*\left[-D_k^{-2} g_k\right]$$

should be

$$\phi_k(d_k) \le \beta_g \phi_k^*[-D_k g_k]$$

or $-D_k^{\frac{1}{2}} g_k = -\hat{g}_k$ since the scaling matrix is different from the one in your original paper?

(4) I am also very confused about the definition of our $\phi_k^*[d_k]$ in page 37 in the thesis, should I keep that $0 \le \alpha \le \beta_k^2$ because in your original paper, $\phi_k^*[d_k]$ has nothing to do with box-bound or break point in our context?

I have sent you the full version of the proof just for your reference in case you need it in the future time. I have tried my best to have corrected every error in it.

I am looking forward to your comments and suggestions.

Best regards,

Zhirong