Question raised during the private PhD defense of Gustavo Quintana-Carapia

Concerning other sorts of prior estimations, e.g., given by explicit function theorems. I wonder if you could get something and apply this bound in order to control the bound of the 2nd order expansion. There exist some good results that could help you devise something a priori which can control the error you are making. Neuberger's results (quite old now) but may be handy for this kind of problem. This can be investigated further.

The investigation of the error bounds for the second order Taylor expansion can be conducted using the continuous Newton method. This is proposed as future research work in Chapter 7.

Neuberger's paper [1] introduces five theorems that are consequence of the continuous Newton method. The Newton method aims to find a zero of the function $F : R^n \to R^n$ by iterating

$$z_{k+1} = z_k - \left(F'(z_k)\right)^{-1} F(z_k) \tag{1}$$

for $k = 0, 1, 2, \ldots$, and for a given initial z_0, assuming that $(F'(y))^{-1}$ exists for some $y \in R^n$. A domain of attraction is the set of all initial values z_0 that lead to a root of F after the convergence $z_0, z_1, z_2, \ldots$. The damped Newton method

$$z_{k+1} = z_k - \delta_k \left(F'(z_k)\right)^{-1} F(z_k) \tag{2}$$

prevents convergence issues, such as chaotic domains of attraction, with $\delta_1, \delta_2, \ldots \in (0, 1)$. The continuous Newton method is a sequence of damped Newton methods. To implement the continuous Newton method, select $T > 0$ and perfom m runs of the damped Newton method with $\delta_k = T/m$, for $k = 1, 2, \ldots, m$. The result z_{k+1} is the zero estimation x_m, and if the sequence of estimates $x_1, x_2, \ldots$ converges, then it converges to the result of the continuous Newton method.

The objective of the continuous Newton method is to find a function $z : [0, \infty) \to R^n$, so that

$$z(0) = x \in R^n, \quad z'(t) = -\left(F'(z(t))\right)^{-1} F(z(t)), \quad t \geq 0, \tag{3}$$

subject to the existence of the limit $u = \lim_{t \to \infty} z(t)$, that satisfies $F(u) = 0$. This implies

$$F'(z(t))z'(t) = -F(z(t)), \tag{4}$$

that is equivalent to

$$(F \circ z)'(t) = -F(z(t)), \tag{5}$$

from where we get

$$F(z(t)) = e^{-t}F(z(0)). \tag{6}$$

Thus, the residual $F(z(t))$ only changes in magnitude and not in direction. This property is essential for the results of the theorems in [1]. In particular, Theorem 5 gives the conditions for the continuous Newton method to ensure finding a root of the function.

Theorem 5 *Suppose that each of H, J, and K is a Banach space, with H compactly embedded in J, that $r > 0$, and that $G : B_r(0) \to K$ is continuous as a function on J. Suppose also that g belongs to K and that for each y in $b_r(0)$ there is an h in $B_r(0)$, where $b_r(0)$ and $B_r(0)$ are open and closed balls in H of radius r centered at 0, such that*

$$\lim_{t \to 0+} -\frac{1}{t}\left(G(y + th) - G(y)\right) = g. \tag{7}$$

Then, there exists u in $B_r(0)$ such that $G(u) = g$.

The theorem implies that u is a zero of the function defined as $F(y) = G(y) - g$. Since G is a continuous function and if $(G'(y))^{-1}$ exists for each y in $B_r(0)$, a metric for the validity of the Theorem (5) hypothesis is the inequality $\left\|(G'(y))^{-1} g\right\| \le r$.

To link the data-driven step-input estimation problem with this theorem, we can take the system of linear equations $\widetilde{\mathbf{y}} = \widetilde{\mathbf{K}}\boldsymbol{\theta}$, then define $G(y) = \widetilde{\mathbf{K}}\boldsymbol{\theta}$, and $g = \widetilde{\mathbf{y}}$, so that we have $F(\boldsymbol{\theta}) = G(\boldsymbol{\theta}) - g$. The derivative of the function $G(\boldsymbol{\theta})$ is given by $G'(\boldsymbol{\theta}) = \widetilde{\mathbf{K}}\mathbf{I}$. Using the pseudo-inverse of $\widetilde{\mathbf{K}}$, the metric $\left\|\widetilde{\mathbf{K}}^\dagger\widetilde{\mathbf{y}}\right\| \le r$ indicates that the estimate $\hat{\boldsymbol{\theta}}$ is found inside a ball of radius r. The questions that remain open are how large can r be?, and how is the radius r related to the perturbation noise variance?

References

[1] J. W. Neuberger, The Continuous Newton's Method, Inverse Functions, and Nash-Moser, The American Mathematical Monthly 114 (2007) 432–437.