Newton's Method and Rate of Convergence.

Inverting $DP(x_n)$ at each step is tantamount to solving a linearized version of the equation $P(x_n)=0$. This is computationally expensive. There is however a gain — faster (here quadratic) convergence.

Suppose Newton's method converges to a solution $x^* \in X$ of the equation $P(x)=0$, and $(DP(x^*))^{-1}$ exists. Furthermore, assume that within an open region containing x^* and the sequence $\{x_n\}$, the quantities $\|D^kP(x)\|$, $k=2,3$ and $\|(DP(x))^{-1}\|$ are bounded.

Then, letting $T(x) = x-(DP(x))^{-1}P(x)$,

$$
\begin{aligned}
x_{n+1} - x^* &= x_n - (DP(x_n))^{-1}P(x_n) - x^* \\
&= x_n - (DP(x_n))^{-1}P(x_n) \\
&\qquad - (x^* - (DP(x^*))^{-1}P(x^*))
\end{aligned}
$$

(by hypothesis $P(x^*)=0$)

$$
\begin{aligned}
&= T(x_n) - T(x^*) \\
DT(x) &= (DP(x^*))^{-1}D^2P(x^*)(DP(x^*))^{-1}P(x^*) \\
&= 0
\end{aligned}
$$

By proposition 2 (Lecture Notes 9(b)), on mean values,

$$\|x_{n+1} - x^*\| \le \frac{1}{2} \sup_{\bar{x}} \|D^2 T(\bar{x})\| \, \|x_n - x^*\|^2$$

where $\bar{x} = x_n + \alpha(x_n - x^*)$, $0 \le \alpha \le 1$.

Hence $\|x_{n+1} - x^*\| \le c \, \|x_n - x^*\|^2$

where $c = \frac{1}{2} \sup_{x \in R} \|D^2 T(x)\|$, a bound depending upon $\|D^3 P(x)\|$.

This is quadratic convergence.

Application 1

Euler Langrange equations in the calculus of variations can be cast in the form

(EL) $\dot{x} = F(x(t))$

$$C x(t_1) = c_1$$

$$D x(t_2) = d_2$$

for problems with fixed end points. Such a boundary value problem can be solved by Newton's method. First (EL) can be written as a

$$x(t) = x(t_1) + \int_{t_1}^{t} F(x(\sigma)) \, d\sigma \qquad t_1 \le t \le t_2.$$

$\leftrightarrow$ $P(x) = 0$ where,

$$(P(x))(t) = x(t) - x(t_1) - \int_{t_1}^{t} F(x(\sigma))\,d\sigma$$

$$t_1 \le t \le t_2 .$$

It is left as an <u>exercise</u> to show that the Fréchet derivative $DP(x)$ is given by

$$\left(DP(x)\cdot h\right)(t) = h(t) - \int_{t_1}^{t} F_x(x(\sigma))\,h(\sigma)\,d\sigma$$

where $F_x(x(\sigma)) = \left[\dfrac{\partial F^i}{\partial x_j}\right]_{x(\sigma)}$.

Newton's method is then equivalent to

$$DP(x_n)\cdot(x_{n+1} - x_n) = -P(x_n)$$

$\Leftrightarrow$

$$x_{n+1}(t) - x_n(t) - \int_{t_1}^{t} F_x(x_n(\sigma))(x_{n+1}(\sigma) - x_n(\sigma))\,d\sigma$$

$$= -x_n(t) + x(t_1) + \int_{t_1}^{t} F(x_n(\sigma))\,d\sigma$$

Differentiating both sides we get,

$$\dot{x}_{n+1}(t) = F_x(x_n(t))\cdot(x_{n+1}(t) - x_n(t)) + F(x_n(t))$$

$$= F_x(x_n(t))\cdot x_{n+1}(t) + \left(F(x_n(t)) - F_x(x_n(t))x_n(t)\right)$$

which is a linear time varying system with input, of the form $\dot{x}_{n+1}(t) = A_n(t)\,x_{n+1}(t) + b_n(t)$

where

$$A_n(t) = F_x(x_n(t))$$

$$b_n(t) = F(x_n(t)) - F_x(x_n(t))\,x_n(t)$$

add to this the boundary conditions

$$C\, x_{n+1}(t_1) = c_1$$

$$D\, x_{n+1}(t_2) = d_2 \,.$$

We can solve this ___linear___ boundary value problem as follows.

Let $\Phi_n(t_2, t)$ be defined by the transition matrix equation

$$\dot{\Phi}_n(t_2, t) = -\Phi_n(t_2, t)\, A_n(t)$$

$$\Phi_n(t_2, t_2) = \mathbb{1} \,.$$

~~let~~ Then

$$x_{n+1}(t_2) = \Phi_n(t_2, t_1)\, x(t_1) + \int_{t_1}^{t_2} \Phi_n(t_2, \sigma)\, b_n(\sigma)\, d\sigma$$

The boundary conditions become

$$C\, x_{n+1}(t_1) = c_1$$

$$D\, \Phi_n(t_2, t_1)\, x_{n+1}(t_1) = d_2 - D \int_{t_1}^{t_2} \Phi_n(t_2, \sigma)\, b_n(\sigma)\, d\sigma$$

solve for $x_{n+1}(t_1)$. It there follows that

$$x_{n+1}(t) = \Phi_n(t, t_1) x(t_1) + \int_{t_1}^{t} \Phi_n(t, \sigma) b_n(\sigma) \, d\sigma$$

where

$\Phi_n(t, t_1)$ satisfies

$$\dot{\Phi}_n(t, t_1) = A_n(t) \, \Phi_n(t, t_1)$$

$$\Phi_n(t_1, t_1) = \mathbb{1} .$$

Of course $\Phi_n(t, t_1) = \Phi_n(t, t_2) \Phi_n(t_2, t_1)$

$$= \Phi_n(t_2, t)^{-1} \Phi_n(t_2, t_1)$$

Thus we solve a linear system at each step of the Newton iteration scheme.

This is a practical method to solve Euler Lagrange equations.