2

الف) روش نیوتن در بهینه سازی برای پیدا کردن min و max تابع با استفاده از تقریب quadratic سری تیلور تابع $f(x)$ بدست می آید.

$$f(x) = \underbrace{f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k)}_{\text{تقریب quadratic}} + \cdots$$

که می توانیم با فرمول زیر x_k را آپدیت کنیم.

$$x_{k+1} = x_k - H^{-1}(x_k)\nabla f(x_k)$$ مانند فرم غیر ماتریسی $$f(x_{k+1}) = f(x_k) - \frac{f'(x_k)}{f''(x_k)}$$

که در این رابطه از یک x_0 شروع کرده و به شرط positive definite بودن H (ماتریس هسین) تا وقتی که ∇f از یک ϵ کمتر شود (تقریبی converge) ادامه می دهیم.

ب) مشکلات روش نیوتن:

در این روش از وارون ماتریس هسین (H^{-1}) استفاده می شود. اگر H وارون پذیر نباشد یا H وجود نداشته باشد نمی توانیم از این روش استفاده کنیم. محاسبه H خود نیز overhead بالایی دارد و ممکن است زمان زیادی ببرد. از طرفی اگر H positive definite هم نباشد ممکن است در saddle point متوقف شویم که min یا max نیست. این الگوریتم وقتی نزدیک به جواب باشیم عالی عمل می کند اما برای شروع و در نقاط دور خوب کار نمی کند.

پ) روش هایی که مشکلات قسمت قبل را حل می کنند به صورت زیر هستند:

روش Quasi-Newton که در این روش به جای ماتریس هسین از تقریب آن استفاده می کنیم.

$$x_{k+1} = x_k - \alpha_k S_k g_k$$

که S_k تقریبی از $H(x_k)$ است.

α_k: stepsize — S_k: ماتریس تخمین هسین — $g_k \rightarrow \nabla f(x_k)$

ت) روش BFGS:

در این روش که ادامه روش Quasi-Newton است داریم:

$$x_{k+1} = x_k - \alpha_k S_k g_k$$

فرمول بدست آوردن S_k در BFGS: $S_{k+1} = S_k + Q_k \quad , \quad S_0 = I$

یا

$$S_{k+1} = \left(I - \frac{\Delta x_k \Delta g_k^T}{\Delta x_k^T \Delta g_k} \right) S_k \left(I - \frac{\Delta x_k \Delta g_k^T}{\Delta x_k^T \Delta g_k} \right) + \frac{\Delta x_k \Delta x_k^T}{\Delta x_k^T \Delta g_k}$$

در این روش می دانیم اگر $S_k > 0$ $\Longleftarrow$ $S_{k+1} > 0$ و همچنین همگرایی تابع در n گام تضمین شده است. به عبارتی $S_k = H_f^{-1}(x^*)$

این روش که توسط ۴ دانشمند پیشنهاد شد می تواند تقریب خوبی برای H در مرحله k به ما بدهد.

روش DFB:

در این روش نیز تقریبی بر H می دهیم! اما نه بر $\cancel{\nabla f}$. داریم:

$$x_{k+1} = x_k - B_k^{-1} \nabla f(x_k)$$

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}$$

$s_k = x_{k+1} - x_k$ و $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$.

با این آپدیت ها B_k ما تقریبی خوب برای ماتریس هسین خواهد بود.