الف) روش نیوتون یک روش عددی برای یافتن مینیمم یا ماکزیمم یک تابع است که از اطلاعات گرادیان و هسین (ماتریس مشتقات مرتبه دوم) استفاده می‌کند. این روش بر پایه تقریب درجه دوم سری تیلور از تابع $f(x)$ در نزدیکی یک نقطه است.

برای یک تابع $f(x)$ که قابل دو بار مشتق‌گیری است، مدل درجه دوم (Approximation) از $f(x)$ در نقطه x_k به صورت زیر است:

$$m_k(x) = f(x_k) + \nabla f(x_k)^T(x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k),$$

که در آن $\nabla f(x_k)$ گرادیان تابع در x_k و $H(x_k)$ ماتریس هسین (ماتریس مشتقات مرتبه دوم) در x_k.

برای مینیمم کردن $m_k(x)$، مشتق آن نسبت به x صفر قرار می‌گیرد:

$$\nabla m_k(x) = \nabla f(x_k) + H(x_k)(x - x_k) = 0.$$

با حل این معادله برای x، رابطه روش نیوتون به دست می‌آید:

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

ب) ۱. محاسبه ماتریس هسین $H(x_k)$ و معکوس آن $H(x_k)^{-1}$ در هر تکرار به خصوص در مسائل با ابعاد بالا، هزینه محاسباتی زیادی دارد ($O(n^3)$). برای مسائل بزرگ، روش نیوتن ممکن است عملی نباشد.

۲. اگر $H(x_k)$ مثبت معین نباشد (یعنی بعضی مقادیر ویژه آن منفی باشند)، روش نیوتون ممکن است به سمت نقطه زینی (Saddle Point) یا حتی نقطه ماکزیمم حرکت کند. در این موارد، روش نیوتن برای مسائل مینیمم‌سازی به درستی کار نمی‌کند.

۳. اگر نقطه شروع x_0 به نقطه بهینه نزدیک نباشد یا گرادیان کوچک نباشد، روش نیوتن ممکن است به کندی همگرا شود یا حتی واگرا شود.

۴. اگر ماتریس هسین $H(x_k)$ منفرد (Singular) یا تقریباً منفرد باشد، محاسبه $H(x_k)^{-1}$ غیرممکن یا ناپایدار می‌شود.

۵. در مسائل غیرکوژ (Non-Convex)، روش نیوتون ممکن است به یک مینیمم محلی یا نقطه زینی همگرا شود و از مینیمم کلی عبور کند.

پ) روش‌های نیوتن تصحیح‌شده:

۱. روش نیوتن با طول گام کاهش‌یافته (Damped Newton Method): در این روش، به جای استفاده از طول گام کامل $\alpha = 1$، از طول گام کوچک‌تری استفاده می‌شود:

$$x_{k+1} = x_k - \alpha_k H(x_k)^{-1}\nabla f(x_k), \quad 0 < \alpha_k \leq 1.$$

مقدار α_k معمولاً با استفاده از جستجوی خط (Line Search) انتخاب می‌شود تا همگرایی تضمین شود. این روش از گام‌های بسیار بزرگ که باعث واگرایی می‌شوند، جلوگیری می‌کند.

۲. روش نیوتن برای مسائل نامعین یا غیرکوژ: اگر ماتریس هسین مثبت معین نباشد، می‌توان از روش‌هایی استفاده کرد که یک ماتریس مثبت معین تخمینی H_k ایجاد کنند:

- تصحیح اسپکترا (Spectral Correction): مقادیر ویژه منفی هسین را به صفر یا یک مقدار مثبت کوچک تغییر می‌دهند.

- افزودن یک مقدار مثبت به قطر هسین (Levenberg-Marquardt):

$$H'_k = H(x_k) + \lambda I, \quad \lambda > 0,$$

که λ یک مقدار مثبت است و I ماتریس همانی است.

این روش از مشکلات ناشی از منفی معین بودن ماتریس هسین جلوگیری می‌کند.

۳. روش نیوتن محدودشده (Trust Region Newton): به جای گام مستقیم $\Delta x_k = -H(x_k)^{-1}\nabla f(x_k)$، مسئله بهینه‌سازی محدود شده حل می‌شود:

$$\min_{\Delta x} \quad m_k(\Delta x) = f(x_k) + \nabla f(x_k)^T \Delta x + \frac{1}{2}\Delta x^T H(x_k)\Delta x,$$

با شرط $\|\Delta x\| \leq \Delta$ (که Δ شعاع منطقه اعتماد است).

این روش از حرکت به نقاط دور از ناحیه معتبر جلوگیری می‌کند و پایداری را افزایش می‌دهد.

ت) روش‌های شبه‌نیوتن تلاش می‌کنند اطلاعات مشتقات مرتبه دوم را با هزینه محاسباتی کمتر و بدون نیاز به محاسبه مستقیم ماتریس هسین، تخمین بزنند. این روش‌ها به جای هسین واقعی، از یک ماتریس تقریب‌زننده B_k استفاده می‌کنند.

روش‌های شبه‌نیوتن یک ماتریس B_{k+1} را تخمین می‌زنند که تقریباً رفتار ماتریس هسین $H(x)$ را مدل‌سازی می‌کند.

شرط به‌روزرسانی اصلی (Secant Equation):

$$B_{k+1}s_k = y_k,$$

که در آن $s_k = x_{k+1} - x_k$ تغییر موقعیت و $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ تغییر گرادیان است.

روش‌های شبه‌نیوتن مسئله زیر را حل می‌کنند:

$$\min_{B} \quad \|B - B_k\|,$$

به طوری که:

$$Bs_k = y_k.$$

اینجا $\|\cdot\|$ معمولاً یک نُرم ماتریسی است (مثلاً نُرم فروبنیوس).

روش **DFP** (Davidon-Fletcher-Powell): در این روش، ماتریس هسین معکوس $H_k = B_k^{-1}$ مستقیماً تخمین زده می‌شود. به‌روزرسانی H_{k+1} به صورت زیر است:

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}.$$

$\frac{s_k s_k^T}{s_k^T y_k}$ این بخش به ماتریس H_k اطلاعات جدید از تغییرات موقعیت و گرادیان اضافه می‌کند.

$\frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}$ این بخش تغییرات گذشته را تنظیم می‌کند تا سازگار با تغییرات جدید باشد.

روش **BFGS** (Broyden-Fletcher-Goldfarb-Shanno): این روش نیز ماتریس هسین معکوس H_{k+1} را به‌روزرسانی می‌کند. به‌روزرسانی به صورت زیر است:

$$H_{k+1} = \left(I - \frac{s_k y_k^T}{y_k^T s_k}\right) H_k \left(I - \frac{y_k s_k^T}{y_k^T s_k}\right) + \frac{s_k s_k^T}{y_k^T s_k}.$$

$I - \frac{s_k y_k^T}{y_k^T s_k}$ این بخش ماتریس فعلی H_k را اصلاح می‌کند.

$\frac{s_k s_k^T}{y_k^T s_k}$ این بخش اطلاعات جدید را اضافه می‌کند.