Lecture FYS4411, February 16, 2024

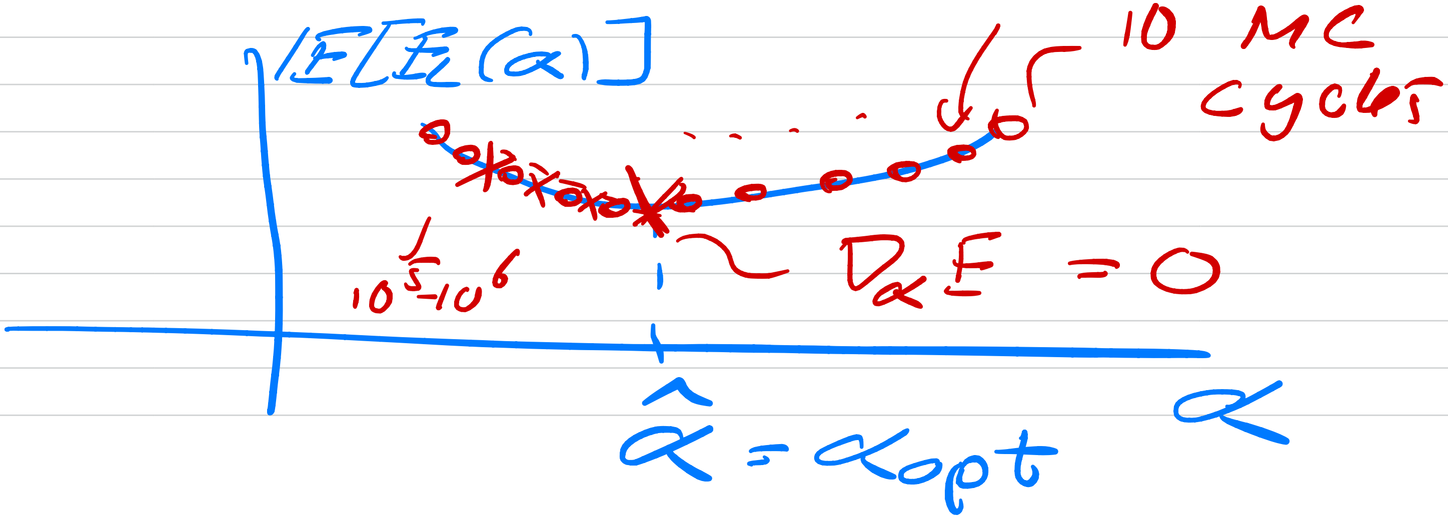
Optimization

$$\mathbb{E}\left[E_L(\vec{\alpha})\right] = \int d\vec{R}\; P_{\vec{\alpha}}(\vec{R})\; E_L(\vec{R};\vec{\alpha})$$

$$P_{\vec{\alpha}}(\vec{R}) = \frac{|\Psi_T(\vec{R};\vec{\alpha})|^2}{\int d\vec{R}\, |\Psi_T(\vec{R};\vec{\alpha})|^2}$$

$$E_L(\vec{R};\vec{\alpha}) = \frac{1}{\Psi_T(\vec{R};\vec{\alpha})} H \Psi_T(\vec{R};\vec{\alpha})$$

$$\hat{\alpha} = \{\alpha\}$$



$$\mathbb{E}[E_L(\vec{\alpha})] = E(\vec{\alpha})$$

$$\vec{\nabla} E(\vec{\alpha}) = g(\vec{\alpha}) \qquad \vec{\alpha} \to \alpha$$

$$= g(\alpha)$$

Taylor expand $E(\alpha)$ around

$$\alpha_{opt} = \hat{\alpha}$$

$$\hat{\alpha} - \alpha_n = b_n \qquad g(\alpha) = g(\alpha_n) = g_n$$

↑ iteration - n -

$$\nabla_\alpha^2 E(\alpha) \to \nabla_\alpha^2 E(\alpha_n) = A_n$$

$$E(\hat{\alpha}) = \underset{= E(\alpha_n)}{E_n} + \underbrace{g_n^T b_n}_{g_n(\hat{\alpha} - \alpha_n)} + \frac{1}{2} b_n^T A_n b_n + O(b_n^3)$$

with $\alpha = \{\alpha_0, \alpha_1\}$

$$A_n = \begin{bmatrix} \left.\frac{\partial^2 E}{\partial \alpha_0^2}\right|_{\alpha=\alpha_n} & \left.\frac{\partial^2 E}{\partial \alpha_0 \partial \alpha_1}\right|_{\alpha=\alpha_n} \\ \left.\frac{\partial^2 E}{\partial \alpha_1 \partial \alpha_0}\right|_{\alpha=\alpha_n} & \left.\frac{\partial^2 E}{\partial \alpha_1^2}\right|_{\alpha=\alpha_n} \end{bmatrix}$$

Neglect $O(b_n^3)$

$$E(\alpha) = \underbrace{E(\alpha_n)}_{E_n} + g_n^T b_n + \frac{1}{2} b_n^T A_n b_n$$

is of the form

$$f(x) = c + g^T x + \frac{1}{2} x^T A x$$

$$\frac{\partial f}{\partial x} = 0 = Ax + g \Rightarrow$$

$$x = A^{-1} g$$

g known

A known

Simple example

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + 10 x_2^2 - 5x_1 - 3x_2$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \frac{1}{2} x^T \overbrace{\begin{bmatrix} 2 & 1 \\ 1 & 20 \end{bmatrix}}^{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \underbrace{\begin{bmatrix} 5 & 3 \end{bmatrix}}_{e^T} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{\partial f}{\partial x} = 0 \Rightarrow x = A^{-1} b$$

$$\frac{\partial f}{\partial x_1} = 0 = 2x_1 + x_2 - 5$$

$$\frac{\partial f}{\partial x_2} = 0 = x_1 + 20x_2 - 3 =$$

guess $x_0 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

what we need is

$\nabla_{\hat{\alpha}} E(\hat{\alpha})$ $\qquad \hat{\alpha} = \alpha_{n+1}$

$\nabla E(\alpha_{n+1}) = 0 = g_n + A_n b_n$

$b_n = \alpha_{n+1} - \alpha_n$

$g_n = \nabla E(\alpha_n) \Rightarrow$

$\alpha_{n+1} - \alpha_n = - A_n^{-1} g_n \Rightarrow$

$$\boxed{\alpha_{n+1} = \alpha_n - A_n^{-1} g_n}$$

Newton-Raphson's method

$x_{n+1} = x_n - f(x_n)/f'(x_n)$

Solve iteratively till

$$\|x_{n+1} - x_n\|_2 \le \varepsilon$$

A_n is a second derivative common to replace A_n with a constant γ_n

How do we find the optimal γ_n?

Expand

$$E(x_n - \gamma_n g_n) = E_n - \gamma_n g_n^T g_n$$

$$x_{n+1} = x_n - \gamma_n g_n$$

$$+ \frac{1}{2} \gamma_n^2 g_n^T A_n g_n + \cancel{O((d_n - \gamma_n g_n)^3)}$$

$$E_{n+1} = E_n - \gamma_n g_n^T g_n + \frac{1}{2} \gamma_n^2 g_n^T A_n g_n$$

depends on γ_{n-1}

$$\frac{dE}{d\gamma_n} = 0 \implies$$

$$\gamma_n = \frac{g_n^T g_n}{g_n^T A_n g_n}$$

(steepest descent)

suppose $A u_n = \lambda u_n$

$u_n^T u_n = 1$

$\gamma_n = \frac{1}{\lambda}$

γ_n as step size has

min value $\frac{1}{\lambda_{max}}$ and

max value $\frac{1}{\lambda_{min}}$

one can show that

$\gamma_n < \frac{2}{\lambda_{max}}$

why do we want to avoid calculating $A_n = \nabla^2 E_n$

$\nabla E_n = \nabla \underline{E}(\alpha)$

Example: 1-dim $\underline{HO}$

$$\psi_T(x;\alpha) = e^{-1/2\,\alpha^2 x^2}$$

$$E(\alpha) = \frac{\int dx\, e^{-1/2\,\alpha^2 x^2} H e^{-1/2\,\alpha^2 x^2}}{\int dx\, e^{-\alpha^2 x^2}}$$

$$= \frac{\int dx\, \psi_\alpha^*(x) H \psi_\alpha(x)}{\int dx\, |\psi_\alpha(x)|^2}$$

$$\frac{dE}{d\alpha} = \frac{2\int dx\, H\, \psi_\alpha(x)\, \frac{d\psi_\alpha(x)}{d\alpha}\, \psi_\alpha(x)}{\int dx\, \psi_\alpha^2(x)}$$

$$- 2\left[\frac{\int dx\, \psi_\alpha^*(x)\, H\, \psi_\alpha(x)}{\int dx\, \psi_\alpha^2(x)}\right] \times$$

$$\left[\frac{\int dx\, \psi_\alpha(x)\, \frac{d\psi_\alpha(x)}{d\alpha}}{\int dx\, \psi_\alpha^2(x)}\right]$$

2nd term :

$$\frac{\int dx \frac{\psi_\alpha}{\psi_\alpha} \psi_\alpha(x) \frac{d\psi_\alpha(x)}{d\alpha}}{\int dx\, |\psi_\alpha|^2}$$

$$= \frac{\int dx\, |\psi_\alpha(x)|^2 \frac{d \ln \psi_\alpha(x)}{d\alpha}}{\int dx\, |\psi_\alpha(x)|^2}$$

$$= \int P_\alpha(x) \frac{d \ln \psi_\alpha(x)}{d\alpha} dx$$

$$\frac{dE}{d\alpha} = 2\left\langle \frac{d\ln\Psi_\alpha}{d\alpha} E_L(\alpha)\right\rangle$$

$$-2\left\langle \frac{d\ln\Psi_\alpha}{d\alpha}\right\rangle \langle E_L(\alpha)\rangle$$

Exercise: show this for the general case

challenge: calculate

$$\left.\frac{d^2E}{d\alpha^2}\right| \frac{dE}{d\alpha} = 0$$

Example Ψ_T in Notebook

$$\Psi_T(\vec{r}_1, \vec{r}_2; \alpha, \beta)$$

$$= \underbrace{e^{-\frac{1}{2}\alpha^2(r_1^2 + r_2^2)}}_{\Psi_{OB}} e^{\frac{r_{12}}{1+\beta r_{12}}}$$

$$r_{12} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$