Quantum Chemistry, lecture notes, 24-Nov-2010, Gerrit C. Groenenboom

I. UNDETERMINED MULTIPLIER METHOD OF LAGRANGE

Let $\boldsymbol{A}$ be a real, regular, symmetric $n \times n$ matrix, and $\boldsymbol{b}$ a vector in $\mathbb{R}^n$. Determine a stationary point of the function $f : \mathbb{R}^n \to \mathbb{R}$

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T \boldsymbol{A}\boldsymbol{x} \tag{1}$$

under the condition that ($g : \mathbb{R}^n \to \mathbb{R}$)

$$g(\boldsymbol{x}) = \boldsymbol{b}^T \boldsymbol{x} - 1 = 0. \tag{2}$$

Solution: according to the undetermined multiplier method of Langrange, we must first solve:

$$\boldsymbol{\nabla} f(\boldsymbol{x}) = \lambda \boldsymbol{\nabla} g(\boldsymbol{x}) \tag{3}$$

Compute the gradient of function g:

$$[\boldsymbol{\nabla} g(\boldsymbol{x})]_k = \frac{\partial}{\partial x_k}\left(\sum_{i=1}^{n} b_i x_i - 1\right) = \sum_{i=1}^{n} b_i \frac{\partial x_i}{\partial x_k} = \sum_{i=1}^{n} b_i \delta_{ik} = b_i. \tag{4}$$

and the gradient of function f

$$[\boldsymbol{\nabla} f(\boldsymbol{x})]_k = \frac{\partial}{\partial k}\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i A_{ij} x_j = \frac{1}{2}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_{ik} A_{ij} x_j + \sum_{i=1}^{n}\sum_{j=1}^{n} x_i A_{ij} \delta_{jk}\right) \tag{5}$$

$$= \frac{1}{2}\left(\sum_{j=1}^{n} A_{kj} x_j + \sum_{i=1}^{n} x_i A_{ik}\right) = \sum_{i=1}^{n} A_{ki} x_i = (\boldsymbol{A}\boldsymbol{x})_k, \tag{6}$$

where we used $A_{kj} = A_{jk}$ ($\boldsymbol{A}$ is symmetric). Hence, Eq. (3) becomes

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{b}. \tag{7}$$

Since $\boldsymbol{A}$ is regular, this set of $n \times n$ linear equations can be solved

$$\boldsymbol{x} = \lambda \boldsymbol{A}^{-1}\boldsymbol{b}. \tag{8}$$

The undetermined multiplier λ is found from the condition

$$g(\boldsymbol{x}) = \boldsymbol{b}^T(\lambda \boldsymbol{A}^{-1}\boldsymbol{b}) - 1 = 0, \tag{9}$$

i.e.,

$$\lambda = (\boldsymbol{b}^T \boldsymbol{A}^{-1}\boldsymbol{b})^{-1} \tag{10}$$

and the stationary point is

$$\boldsymbol{x} = (\boldsymbol{b}^T \boldsymbol{A}^{-1}\boldsymbol{b})^{-1}\boldsymbol{A}^{-1}\boldsymbol{b}. \tag{11}$$

II. METHOD OF LAGRANGE: COMPLEX VARIABLES

Let $\boldsymbol{A}$ be a regular, Hermitian $n \times n$ matrix, and $\boldsymbol{b}$ a vector in $\mathbb{C}^n$. Determine a stationary point of the function $f : \mathbb{C}^n \to \mathbb{R}$

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\dagger \boldsymbol{A}\boldsymbol{x} \tag{12}$$

under the condition that ($g : \mathbb{C}^n \to \mathbb{R}$):

$$g(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{b}^\dagger \boldsymbol{x} + \boldsymbol{x}^\dagger \boldsymbol{b}) - 1 = 0. \tag{13}$$

We follow the same procedure as before, except that we treat $\{x_1, x_2, \ldots, x_n, x_1^*, x_2^*, \ldots, x_n^*\}$ as $2n$ independent variables. In particular, we use for all $i = 1, \ldots, n$ and $j = 1, \ldots, n$

$$\frac{\partial x_i}{\partial x_j^*} = \frac{\partial x_i^*}{\partial x_j} = 0, \tag{14}$$

even though for $i = j$ these derivatives do not exist for complex variables. For the gradients we find

$$[\boldsymbol{\nabla} g(\boldsymbol{x})]_k = \frac{\partial}{\partial x_k} \frac{1}{2} \left(\sum_i b_i^* x_i + \sum_i x_i^* b_i - 1 \right) = \frac{1}{2} b_k^* \tag{15}$$

and

$$[\boldsymbol{\nabla}^* g(\boldsymbol{x})]_k = \frac{\partial}{\partial x_k^*} \frac{1}{2} \left(\sum_i b_i^* x_i + \sum_i x_i^* b_i - 1 \right) = \frac{1}{2} b_k. \tag{16}$$

and

$$[\boldsymbol{\nabla} f(\boldsymbol{x})]_k = \frac{\partial}{\partial x_k} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n x_i^* A_{ij} x_j = \frac{1}{2} \sum_{i=1}^n x_i^* A_{ik} = \frac{1}{2} (\boldsymbol{x}^\dagger \boldsymbol{A})_k \tag{17}$$

and

$$[\boldsymbol{\nabla}^* f(\boldsymbol{x})]_k = \frac{\partial}{\partial x_k^*} \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n x_i^* A_{ij} x_j = \frac{1}{2} \sum_{j=1}^n A_{kj} x_j = \frac{1}{2} (\boldsymbol{A}\boldsymbol{x})_k \tag{18}$$

Equation (3) now becomes two equations

$$\frac{1}{2} \boldsymbol{x}^\dagger \boldsymbol{A} = \lambda \frac{1}{2} \boldsymbol{b}^\dagger \tag{19}$$

and

$$\frac{1}{2} \boldsymbol{A}\boldsymbol{x} = \lambda \frac{1}{2} \boldsymbol{b} \tag{20}$$

These two equations are the same for real λ since $\boldsymbol{A}^\dagger = \boldsymbol{A}$. The solution is

$$\boldsymbol{x} = \lambda \boldsymbol{A}^{-1} \boldsymbol{b} \tag{21}$$

The condition gives

$$\frac{1}{2} [\boldsymbol{b}^\dagger \lambda \boldsymbol{A}^{-1} \boldsymbol{b} + (\lambda \boldsymbol{A}^{-1} \boldsymbol{b})^\dagger \boldsymbol{b}] = 1 \tag{22}$$

and with $(\lambda \boldsymbol{A}^{-1} \boldsymbol{b})^\dagger = \lambda \boldsymbol{b}^\dagger \boldsymbol{A}^{-1}$ (λ is real and $\boldsymbol{A}$ is Hermitian) we get

$$\lambda = (\boldsymbol{b}^\dagger \boldsymbol{A}^{-1} \boldsymbol{b})^{-1} \tag{23}$$

and finally

$$\boldsymbol{x} = (\boldsymbol{b}^\dagger \boldsymbol{A}^{-1} \boldsymbol{b})^{-1} \boldsymbol{A} \boldsymbol{b} \tag{24}$$

Notice that since the functions f and g are real, it is sufficient to do only half the work, e.g. only work out the $\boldsymbol{\nabla}^*$ part, Eqs. (16), (18), and (20).

Exercise

The expectation value of the Hamiltonian for a wave function expanded in an n-dimensional orthonormal basis may be written as

$$E = \frac{\boldsymbol{c}^\dagger \boldsymbol{H} \boldsymbol{c}}{\boldsymbol{c}^\dagger \boldsymbol{c}} \tag{25}$$

where $\boldsymbol{c} \in \mathbb{C}^n$, and $\boldsymbol{H}$ is the $n \times n$ Hermitian Hamiltonian matrix. Use the undetermined multiplier method of Lagrange to minimize

$$f(\boldsymbol{c}) = \boldsymbol{c}^\dagger \boldsymbol{H} \boldsymbol{c} \tag{26}$$

under the condition that the wave function is normalized

$$g(\boldsymbol{c}) = \boldsymbol{c}^\dagger \boldsymbol{c} - 1 = 0. \tag{27}$$