Majorizing a Multivariate Polynomial over the Unit Sphere

Patrick J.F. Groenen* Jan de Leeuw†

June 27, 2011

*Econometric Institute, Erasmus University Rotterdam, The Netherlands, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands (e-mail: groenen@few.eur.nl).

†Statistics, UCLA (e-mail: deleeuw@stat.ucla.edu).

Abstract

Abstract

Minimizing a multivariate polynomial over the unit sphere can be quadratically majorized. By using Gerschgorin disks for the interval of the eigenvalues a reasonably sharp upper bound of the largest eigenvalue of the Hessian can be obtained.

Keywords: .

1 Introduction

Let $P(x) = \sum_\alpha p_\alpha x^\alpha$ be a multivariate polynomial with α are vectors of m nonnegative integers and

$$x^\alpha = \prod_{j=1}^{m} x_j^{\alpha_j}. \tag{1}$$

The topic of this paper is to minimize $P(x)$ subject to $x'x = 1$. The gradient of $P(x)$ equals

$$g(x) = \frac{\partial P(x)}{\partial x} = \sum_\alpha p_\alpha x^\alpha z \text{ with } z_j = \begin{cases} \alpha_j / x_j & \text{if } \alpha_j \geq 1 \\ 0 & \text{if } \alpha_j = 0 \end{cases}. \tag{2}$$

The Hessian H of $P(x)$ equals

$$H(x) = \frac{\partial^2 P(x)}{\partial x \partial x'} = \sum_\alpha p_\alpha x^\alpha (zz' - D)$$

with D a diagonal matrix with elements $d_{jj} =$

$$d_{jj} = \begin{cases} \alpha_j / x_j^2 & \text{if } \alpha_j \geq 2 \\ 0 & \text{if } \alpha_j < 2 \end{cases}.$$

Let H_α be the Hessian of $p_\alpha x^\alpha = p_\alpha \prod_{j=1}^m x_j^{\alpha_j}$. Then, the element $h_{ij|\alpha}$ equals

$$h_{ij|\alpha} = \begin{cases} p_\alpha w_{ij|\alpha} \prod_{\ell=1}^m x_\ell^{\beta_\ell} & \text{if } j \neq i \\ p_\alpha w_{ii|\alpha} \prod_{\ell=1}^m x_\ell^{\gamma_\ell} & \text{if } j = i \end{cases}.$$

with

$$\begin{aligned} \beta_\ell &= \begin{cases} \alpha_\ell & \text{if } \ell \neq i \text{ and } \ell \neq j \\ \alpha_\ell - 1 & \text{if } \ell = i \text{ and } \alpha_\ell \geq 1 \\ \alpha_\ell - 1 & \text{if } \ell = j \text{ and } \alpha_\ell \geq 1 \\ 0 & \text{if } \alpha_\ell = 0 \end{cases} \\ \gamma_\ell &= \begin{cases} \alpha_\ell & \text{if } \ell \neq i \\ \alpha_\ell - 2 & \text{if } \ell = i \text{ and } \alpha_\ell \geq 2 \\ 0 & \text{if } \ell = i \text{ and } \alpha_\ell \leq 1 \end{cases} \\ w_{ij|\alpha} &= \begin{cases} \alpha_i \alpha_j & \text{if } j \neq i \\ \alpha_i(\alpha_i - 1) & \text{if } j = i \end{cases} \end{aligned}$$

Obviously, $H = \sum_\alpha H_\alpha$.

For quadratic majorization subject to the length constraint, we need an estimate of the largest eigenvalue of the Hessian H for any admissible x.

Note that H is square and symmetric so that its eigenvalues are real. The Gerschgorin disk theorem gives rowwise bounds on the size of the (unsorted) eigenvalues λ_i, that is,

$$h_{ii} - \sum_{j \neq i} |h_{ij}| \leq \lambda_i \leq h_{ii} + \sum_{j \neq i} |h_{ij}|. \tag{3}$$

Therefore, we need to determine the largest value that each $|h_{ij}|$ can take knowing that $x'x = 1$. Note that $|h_{ij}| = \sum_\alpha |h_{ij|\alpha}|$.

We can use the weighted geometric-arithmetic mean inequality on each H_α that says

$$\begin{aligned} \left(\prod_{j=1}^{m} |x_j|^{\beta_j} \right)^{1/m} &\leq \frac{\sum_{j=1}^{m} \beta_j |x_j|}{\sum_{j=1}^{m} \beta_j} \\ \prod_{j=1}^{m} |x_j|^{\beta_j} &\leq \left(\frac{\sum_{j=1}^{m} \beta_j |x_j|}{\sum_{j=1}^{m} \beta_j} \right)^m . \end{aligned}$$

If all $\beta_j > 0$ then the upper bound is attained whenever all x_j are equal. Obviously, the x_js with $\beta_j = 0$ do not contribute and for attaining an upper bound should be chosen as zero. Let m_{pos} be the number of nonzero β_j. Combining this with the restriction that the the sum of the x_j^2s with nonzero β_j must equal 1 gives $x_j = m_{\text{pos}}^{-1/2}$ if $\beta_j \neq 0$ and $x_j = 0$ if $\beta_j = 0$. Consequently,

$$\prod_{j=1}^{m} |x_j|^{\beta_j} \leq m_{\text{pos}}^{-m_{\text{pos}}/2}.$$

As we are only interested in the largest absolute value of h_{ij}, we can remove the absolute value signs on the left hand side of the geometric-arithmetic mean, that is,

$$\prod_{j=1}^{m} x_j^{\beta_j} \leq m_{\text{pos}}^{-m_{\text{pos}}/2}.$$

and multiplying the left hand side by $p_\alpha w_{ij|\alpha}$ and the right hand side by $|p_\alpha|$ retains the upper bound for the inequality as

$$p_\alpha w_{ij|\alpha} \prod_{j=1}^{m} x_j^{\beta_j} \leq |p_\alpha| w_{ij|\alpha} m_{\text{pos}}^{-m_{\text{pos}}/2}.$$

Note that each $h_{ij|\alpha}$ can be expressed as $p_\alpha w_{ij|\alpha} \prod_{j=1}^{m} x_j^{\beta_j}$, thus also for each $|h_{ij|\alpha}|$.

2 Example

Consider the polynomial

$$P(x_1, x_2) = 3 + 4x_1 - 5x_2 + x_1^2 x_2^2 - 4x_1^3 \tag{4}$$

so that

$$\begin{aligned} P(x_1, x_2) &= 3 + 4x_1 - 5x_2 + x_1^2 x_2^2 - 4x_1^3 = \sum_{k=1}^{5} P_k(x_1, x_2) && (5) \\ & && (6) \end{aligned}$$

with

$$\begin{aligned} P_1(x_1, x_2) &= 3 && (7) \\ P_2(x_1, x_2) &= 4x_1 && (8) \\ P_3(x_1, x_2) &= -5x_2 && (9) \\ P_4(x_1, x_2) &= x_1^2 x_2^2 && (10) \\ P_5(x_1, x_2) &= -4x_1^3. && (11) \end{aligned}$$

The gradients of $P_k(x_1, x_2)$ are

$$\begin{aligned} g_1(x_1, x_2) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (12) \\ g_2(x_1, x_2) &= \begin{bmatrix} 4 \\ 0 \end{bmatrix} && (13) \\ g_3(x_1, x_2) &= \begin{bmatrix} 0 \\ -5 \end{bmatrix} && (14) \\ g_4(x_1, x_2) &= \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} && (15) \\ g_5(x_1, x_2) &= \begin{bmatrix} -12x_1^2 \\ 0 \end{bmatrix} && (16) \\ & && (17) \end{aligned}$$

and Hessian are

$$\begin{aligned} H_1(x_1, x_2) = H_2(x_1, x_2) &= H_3(x_1, x_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} && (18) \\ H_4(x_1, x_2) &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} && (19) \\ H_5(x_1, x_2) &= \begin{bmatrix} -24x_1 & 0 \\ 0 & 0 \end{bmatrix} && (20) \\ & && (21) \end{aligned}$$

We can either use m or m_{pos}. Using m gives an upper bound of $2 + |-24|2^{-2/2} = 14$ as the largest eigenvalue. Using m_{pos} gives $2+|-24|1^{-1/2} = 26$.