

On the Criteria for Receiving a Research Productivity Fellowship from the Brazilian National Council for Scientific and Technological Development in Mathematics

M.E. VALLE^{1*} and F. SAKURAY²

Received on October 8, 2012 / Accepted on November 6, 2014

ABSTRACT. The fellowship of research productivity (PQ) granted by the national council for scientific and technological development (CNPq), besides the financial support, renders a significant status among Brazilian researchers of all areas of knowledge. Consequently, both the profile and the criteria for holding a PQ fellowship become of interest to the entire Brazilian scientific community. In this paper, we model the decision criteria as a weighted sum of the scientific production and the supervisory experience of an applicant for PQ fellowship. The scientific production is measured as the number of publications grouped according to the QUALIS system provided by the Brazilian federal agency for the improvement of higher education (CAPES). The Lattes curricula of PQ fellows in the field of mathematics, along with the curricula of many non-PQ researchers of similar institutions, were used to estimate a criteria for receiving PQ fellowship in category 2. By allowing a certain tolerance, the model reproduced the decision criteria within acceptable bounds over a database composed of 320 curricula.

Keywords: scientific production, research productivity fellowship, mathematics, quadratic programming problem..

1 INTRODUCTION

The *national council for scientific and technological development* (CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, in Portuguese), linked to the Brazilian *ministry of science and technology*, is one of the major public institution for the support of science, technology, and innovation in Brazil. Besides funds for research projects, grants for the purchasing of equipment, and programs for the production of scientific knowledge and the establishment of research networks, CNPq also offers fellowships in Brazil and abroad. In particular, the so-called *fellowship of research productivity* (PQ) have been conceived in the 1970s as a way to encourage

*Corresponding author: Marcos Eduardo Valle

¹Department of Applied Mathematics, University of Campinas, 13083-859 Campinas, SP, Brazil.
E-mails: valle@ime.unicamp.br; mevalle@gmail.com

²Department of Computer Science, State University of Londrina, 86051-990 Londrina, PR, Brazil.
E-mails: sakuray@uel.br; fabio.sakuray@gmail.com

researchers with outstanding scientific production in their fields. Currently, Brazilian researchers of all areas of knowledge desire the PQ fellowship due to the status that comes with it [1, 7, 8]. As a consequence, both the profile and the criteria for holding a PQ fellowship become of interest to the entire scientific community.

The profile and the scientific production of research productivity fellows in various areas of knowledge, including communication science [11], medicine [2, 7, 6, 9], chemistry [8], and zoology [10], have been investigated in the last years. However, as far as we know, there is no study concerning the profile and the production of PQ fellows in the field of mathematics (including applied mathematics). Also, we noted that most papers in the literature provide demographic or statistical information – such as the mean scientific production or the distribution based on gender, region, or institution – of the research productivity fellows. In contrast, this paper provides a quantitative study on the criteria for receiving a PQ fellowship. Specifically, we provide a rule which helps to decide whether a researcher applicant is worthy to receive a PQ fellowship in mathematics.

First of all, the fellowship of research productivity is divided in two categories, called PQ1 and PQ2. The first category is subdivided in four levels, which are referred to as A, B, C, and D, while the category PQ2 has no subdivisions. Non-fellow applicants are eligible only for the second category. Therefore, we will focus only on the criteria for receiving a PQ2 fellowship.

Accordingly to Appendix I of the document that specifies the rules for individual fellowship of CNPq, the resolution RN-016/2006³, the criteria for receiving a PQ2 fellowship must comprise, besides the evaluation of a research project by ad-hoc referees, the scientific production as well as the development of human resources in the last five years. Evidently, the scientific production should be relevant to the field of mathematics. In order to measure the relevance of a publication, we classified the scientific production of a researcher according to the QUALIS system provided by the *Brazilian federal agency for the improvement of higher education* (CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, in Portuguese).

The QUALIS system have been created to assess the scientific production of graduate programs in Brazilian institutions and it is used, in part, to manage the distribution of financial resources. In general terms, journals are ranked into the classes A1, A2, B1, B2, B3, B4, B5, and C, where A1 and C correspond respectively to the higher and lower ranks in the QUALIS system⁴. Moreover, the QUALIS is based on information collected by CAPES, which is concerned with the scientific production of graduate programs. As a consequence, a journal that have not published an article linked to the scientific production of a graduate program may not be ranked in the QUALIS yet. In this case, we attributed the label “N”, which refers to a non-classified journal.

The current QUALIS comprehend the scientific production of graduate programs from 2010 to 2012. The raking of journals in the disciplines of mathematics and statistics is mainly based on

³Website of CNPq. Available at: <http://www.cnpq.br>. Accessed on October, 2014.

⁴Website of CAPES. Available at <http://www.capes.gov.br/avaliacao/instrumentos-de-apoio/classificacao-da-producao-intelectual>, Accessed on October, 2014.

the *impact factor*, the *cited half-live*, and the *article influence score* (AIS) [4]. Also, subjective information such as the editorial board, aims and scope of the journal, and samples of articles has been adopted in case of nonexistent or inadequate indexes.

In sum, let us suppose that we can attribute to each applicant a score based on the scientific production and the formation of human resources since 2009. The scientific production is measured according to the number of papers published in each class of the QUALIS in mathematics and statistics, including the non-classified group. The formation of human resources is measured according to the number of masters and doctors which received their degrees under the supervision of the applicant. All the necessary information can be obtained from the Lattes curriculum of a researcher, which is publicly available at the Lattes platform maintained by CNPq.

In mathematical terms, the score of a researcher is given by the following weighted sum

$$\begin{aligned}
 S(\mathbf{x}) = & \alpha_{A1}x_{A1} + \alpha_{A2}x_{A2} + \alpha_{B1}x_{B1} + \alpha_{B2}x_{B2} + \alpha_{B3}x_{B3} + \alpha_{B4}x_{B4} \\
 & + \alpha_{B5}x_{B5} + \alpha_Cx_C + \alpha_Nx_N + \alpha_Mx_M + \alpha_Dx_D,
 \end{aligned} \tag{1.1}$$

where $x_{A1}, x_{A2}, x_{B1}, x_{B2}, x_{B3}, x_{B4}, x_{B5}, x_C$, and x_N denote respectively the number of papers published in journals classified as $A1, A2, B1, B2, B3, B4, B5, C$, and journals not listed in the QUALIS system. The values x_M and x_D correspond respectively to the number of masters and doctors supervised by the applicant. Alternatively, (1.1) can be written compactly as follows where $\mathbf{x} = [x_{A1}, x_{A2}, \dots, x_D]^T$ and $\boldsymbol{\alpha} = [\alpha_{A1}, \alpha_{A2}, \dots, \alpha_D]^T$ are column vectors in \mathbb{R}^{11} .

$$S(\mathbf{x}) = \boldsymbol{\alpha}^T \mathbf{x}. \tag{1.2}$$

Note that the score is based only on the scientific production and the formation of human resources in the last five years. It does not take into account any information on gender, region, or institution of an applicant. Furthermore, since the classes $A1$ and C correspond respectively to the higher and lower ranks in the QUALIS system, we assume that $\alpha_{A1} \geq \alpha_{A2} \geq \alpha_{B1} \geq \alpha_{B2} \geq \alpha_{B3} \geq \alpha_{B4} \geq \alpha_{B5} \geq \alpha_C \geq 0$. We also impose the inequality $\alpha_D \geq \alpha_M$, which means that a doctorate supervision cannot weight less than a master supervision. However, no restriction have been imposed on the coefficient α_N except that it is a non-negative real number, i.e., $\alpha_N \geq 0$.

The score $S(\mathbf{x})$ yields, in some sense, a quantitative measure of the worthiness of an applicant \mathbf{x} to receive a PQ2 fellowship in mathematics. Specifically, an objective decision criteria can be formulated as follows

$$\begin{cases} \text{The applicant } \mathbf{x} \text{ receives a PQ2 fellowship if } S(\mathbf{x}) \geq \theta, \\ \text{The applicant } \mathbf{x} \text{ does not receive a PQ2 fellowship if } S(\mathbf{x}) < \theta, \end{cases} \tag{1.3}$$

where $\theta \geq 0$ denotes the threshold for receiving a PQ2 fellowship. Note that the objective criteria given by (1.3) does not change if both $\boldsymbol{\alpha}$ and θ are multiplied by a constant $\kappa > 0$. In particular, it is convenient to multiply both sides of the inequalities in (1.3) by $\kappa = 100/\theta$, given that $\theta > 0$. This corresponds to arbitrarily choose $\theta = 100$ in (1.3).

The next section discusses the methodology used to estimate the weight vector α and the threshold θ . An estimation of the decision criteria for receiving a PQ2 fellowship as well as a brief discussion of the results is given in Section 3. Some concluding remarks are given in Section 4. This final section also provides a soft version of the decision criteria given by (1.3). The paper finishes presenting briefly an alternative methodology for estimating the weight vector α and the threshold θ in Appendix A.

2 MATERIAL AND METHODS

First of all, we elaborated a list with the names of the researchers that received a PQ2 fellowship in 2013 and 2014. The names of the PQ2 fellows have been collected from CNPq website⁵. In addition, for each name in the list, we collected the Lattes curricula of 2 members – preferably from the same department – that do not hold a fellowship of research productivity. The Lattes curriculum of all names in our list have been collected during September 2014.

We would like to point out that, in order to avoid a biased experiment, the non-PQ researchers have been selected without any specific criteria by an undergraduate student. Indeed, our student collected the curricula of 34 non-PQ researchers with no scientific production since 2009. Because a non-productive academic will hardly apply for a PQ2 fellowship, we removed these Lattes curricula from our database.

Summarizing, we elaborated a database comprising 320 Lattes curricula, in which 118 correspond to successful applicants for PQ2 fellowship in 2013 and 2014. The remaining 202 curricula represent non-PQ researchers with at least one scientific paper published since 2009. This database have been processed by a software implemented in the `python` programming language. Briefly, the software gathers both the scientific production and the formation of human resources from a Lattes curriculum. The scientific production is ranked using the QUALIS system. In addition, the software adds a label for the successful applicants.

In mathematical terms, the software yielded a data set $\mathcal{D} = \{(\mathbf{x}_1, d_1), \dots, (\mathbf{x}_{320}, d_{320})\}$, where $\mathbf{x}_\xi = [x_{\xi,A1}, \dots, x_{\xi,D}]^T \in \mathbb{R}^{11}$ and $d_\xi \in \{-1, +1\}$ for all $\xi = 1, \dots, 320$. The components of \mathbf{x}_ξ correspond respectively to the numbers of publications sorted according to the groups $A1, A2, B1, B2, B3, B4, B5, C$, and N , followed by the number of masters and doctors under the supervision of the researcher \mathbf{x}_ξ . We have $d_\xi = +1$ if \mathbf{x}_ξ corresponds to a PQ2 fellow and $d_\xi = -1$ if \mathbf{x}_ξ represents a non-PQ researcher.

⁵Available at: <http://www.cnpq.br/web/guest/bolsistas-vigentes> by choosing in the link “fellowship, current fellowships” the options: “Brasil” (country – Brazil), “Todos os Estados” (all provinces), “Todas as Instituições” (all institutions), “Ciências Exatas e da Terra” (area – exact and earth sciences), “Matemática” (sub-area – mathematics), “Produtividade em Pesquisa” (kind of fellowship – research productivity), and “2” (category – 2).

In principle, a decision criteria based on (1.2) and (1.3) should satisfy the following inequalities for all $\xi = 1, 2, \dots, 320$ besides the constraints $\alpha_{A1} \geq \alpha_{A2} \geq \alpha_{B1} \geq \alpha_{B2} \geq \alpha_{B3} \geq \alpha_{B4} \geq \alpha_{B5} \geq \alpha_C \geq 0, \alpha_N \geq 0, \alpha_D \geq \alpha_M \geq 0$, and $\theta \geq 0$:

$$\begin{cases} \alpha^T \mathbf{x}_\xi - \theta \geq +\rho & \text{if } d_\xi = +1, \\ \alpha^T \mathbf{x}_\xi - \theta \leq -\rho & \text{if } d_\xi = -1, \end{cases} \tag{2.1}$$

where $\rho > 0$ denotes the margin of separation between the two classes: the class of PQ fellows and the class of non-PQ researchers. Furthermore, the best decision criteria is given by the weight vector α and the threshold θ that maximize ρ .

We would like to remark that the two inequalities in (2.1) can be combined as follows

$$d_\xi (\mathbf{w}^T \mathbf{x}_\xi - b) \geq 1, \quad \forall \xi = 1, \dots, 320, \tag{2.2}$$

where $\mathbf{w} = \alpha/\rho$ and $b = \theta/\rho$. Also, the following arguments point out that maximizing the margin of separation ρ is equivalent to minimizing the Euclidean norm $\|\mathbf{w}\|$ of the weight vector $\mathbf{w} = [w_1, \dots, w_{11}]^T \in \mathbb{R}^{11}$. For details, we invite the reader to consult [5, 12, 13].

Suppose that the weight vector \mathbf{w} and the threshold b satisfy the constraints (2.2). Also, let $\pi = \{\mathbf{x} \in \mathbb{R}^{11} : \mathbf{w}^T \mathbf{x} - b = 0\}$ denote the corresponding affine hyperplane that separates PQ from non-PQ fellows. Since $\mathbf{w}/\|\mathbf{w}\|$ is a unit vector orthogonal to π , any vector \mathbf{x} can be expressed as

$$\mathbf{x} = \mathbf{x}_\pi + r(\mathbf{x}) \frac{\mathbf{w}}{\|\mathbf{w}\|}, \tag{2.3}$$

where \mathbf{x}_π is the normal projection of \mathbf{x} onto π . Furthermore, $r(\mathbf{x})$ measures the algebraic distance from \mathbf{x} to π , that is, $r(\mathbf{x})$ is positive if \mathbf{x} is on one side of π and negative if \mathbf{x} is on the other side. Alternatively, multiplying by \mathbf{w}^T and subtracting b to both sides of (2.3), we obtain

$$r(\mathbf{x}) = \frac{\mathbf{w}^T \mathbf{x} - b}{\|\mathbf{w}\|}, \quad \forall \mathbf{x} \in \mathbb{R}^{11}. \tag{2.4}$$

Now, since \mathbf{w} and b satisfy (2.2), we conclude that the distance between the vector \mathbf{x}_ξ and the affine hyperplane π satisfy the inequality $|r(\mathbf{x}_\xi)| \geq 1/\|\mathbf{w}\|$ for all $\xi = 1, \dots, 320$. Therefore, maximizing r , which is half the margin of separation between PQ and non-PQ fellows, is equivalent to minimizing the Euclidean norm of \mathbf{w} under the constraints (2.2).

Concluding, in theory, the optimal decision criteria based on (1.2) and (1.3) can be obtained by solving the quadratic programming problem:

$$\begin{cases} \underset{\mathbf{w}}{\text{minimize}} & \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{subject to} & d_\xi (\mathbf{w}^T \mathbf{x}_\xi - b) \geq 1, \quad \forall \xi = 1, \dots, 320, \\ & w_1 \geq w_2 \geq w_3 \geq w_4 \geq w_5 \geq w_6 \geq w_7 \geq w_8, w_{11} \geq w_{10}, \\ & \mathbf{w} = [w_1, \dots, w_{11}]^T \in \mathbb{R}_+^{11}, \text{ and } b \in \mathbb{R}_+, \end{cases} \tag{2.5}$$

where \mathbb{R}_+ denotes the set of non-negative real numbers. In practice, however, the quadratic problem (2.5) does not admit a solution given the Lattes curricula database. Precisely, no weight vector \mathbf{w} and threshold b satisfy all the constraints in (2.5). This remark confirms that the decision criteria given by (1.2) and (1.3) have not been strictly adopted by the administrative council for the mathematical sciences at CNPq. Indeed, we have not considered the merit of the research project that must be submitted by a PQ fellowship applicant. Here, we implicitly assumed that the research project of all applicants have been approved by the ad-hoc referees. Furthermore, on one hand, the administrative council for the mathematical sciences takes into account the regularity of the scientific production. Also, young researchers with potential for leadership are considered by this council. On the other hand, the quantitative model given by (1.2) and (1.3) does not take into account these subjective or hard to measure criteria.

In view of the remarks in the preceding paragraph, we replaced the quadratic programming problem given by (2.5) by a bilevel optimization problem, that is, two hierarchical optimization problems in which the solution of the first defines a constraint of the second optimization problem [3]. An alternative approach which yielded similar results is briefly described in Appendix A.

The first optimization problem minimizes the classification error given the data set \mathcal{D} . Specifically, let us introduce non-negative variables s_1, s_2, \dots, s_{320} , called *slack variables* in the literature, and replace the constraints (2.2) by

$$d_\xi (\mathbf{w}^T \mathbf{x}_\xi - b) \geq 1 - s_\xi, \quad \text{with } s_\xi \geq 0, \quad \text{for all } \xi = 1, \dots, 320. \quad (2.6)$$

Now, let us define the classification error as the sum of the slack variables, that is, the classification error is given by $\mathcal{E} = \sum_{\xi=1}^{320} s_\xi$. Therefore, the first optimization problem is formulated as the following linear programming problem:

$$\left\{ \begin{array}{l} \text{minimize}_{\mathbf{s}} \quad \sum_{\xi=1}^{320} s_\xi \\ \text{subject to} \quad d_\xi (\mathbf{w}^T \mathbf{x}_\xi - b) \geq 1 - s_\xi, \quad \forall \xi = 1, \dots, 320, \\ \quad \quad \quad w_1 \geq w_2 \geq w_3 \geq w_4 \geq w_5 \geq w_6 \geq w_7 \geq w_8, w_{11} \geq w_{10}, \\ \quad \quad \quad \mathbf{w} \in \mathbb{R}_+^{11}, b \in \mathbb{R}_+, \text{ and } \mathbf{s} \in \mathbb{R}_+^{320}, \end{array} \right. \quad (2.7)$$

where $\mathbf{s} = [s_1, s_2, \dots, s_{320}]^T$ is the vector whose components are the slack variables. We solved (2.7) using the GNU GLPK library for GNU Octave⁶, which converged to a solution. The minimum value of the classification error obtained by the linear programming solver was $\mathcal{E}^* = 156.20$.

The second optimization problem determines a weight vector \mathbf{w} and a threshold b for which the margin of separation between PQ and non-PQ fellows is maximized but the classification

⁶Available at: <http://www.gnu.org/software/octave/>. Accessed on October, 2014.

error does not exceed \mathcal{E}^* . In mathematical terms, we have the following quadratic programming problem

$$\left\{ \begin{array}{l} \underset{\mathbf{w}}{\text{minimize}} \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{subject to} \quad \sum_{\xi=1}^{320} s_{\xi} \leq \mathcal{E}^*, \\ d_{\xi}(\mathbf{w}^T \mathbf{x}_{\xi} - b) \geq 1 - s_{\xi}, \quad \forall \xi = 1, \dots, 320, \\ w_1 \geq w_2 \geq w_3 \geq w_4 \geq w_5 \geq w_6 \geq w_7 \geq w_8, w_{11} \geq w_{10}, \\ \mathbf{w} \in \mathbb{R}_+^{11}, b \in \mathbb{R}_+, \text{ and } \mathbf{s} \in \mathbb{R}_+^{320}, \end{array} \right. \quad (2.8)$$

which have been solved using the *quadratic programming solver* for GNU Octave. The GNU Octave solver converged to

$$\mathbf{w}^* = [0.69, 0.42, 0.06, 0.06, 0.06, 0.01, 0.00, 0.00, 0.03, 0.00, 0.03]^T \text{ and } b^* = 1.83, \quad (2.9)$$

for which the sum of the slack variables is 156.20 yet. Finally, in order to simplify the interpretation of the results, we defined $\theta = 100$ and $\alpha = \theta \mathbf{w}^* / b^*$, where \mathbf{w}^* and b^* denote the optimal solutions given by (2.9).

3 RESULTS AND DISCUSSION

The bilevel optimization problem defined by (2.7) and (2.8) yielded the threshold $\theta = 100$ and the weight vector

$$\alpha = [\underbrace{37.87}_{A1}, \underbrace{22.74}_{A2}, \underbrace{3.03}_{B1}, \underbrace{3.03}_{B2}, \underbrace{3.03}_{B3}, \underbrace{0.33}_{B4}, \underbrace{0.0}_{B5}, \underbrace{0.0}_{C}, \underbrace{1.52}_{N}, \underbrace{0.0}_{M}, \underbrace{1.55}_{D}]^T, \quad (3.1)$$

as an estimation for an applicant to be contemplated with a PQ2 fellowship in mathematics.

Note that publications in periodicals ranked B , C , or not listed in the QUALIS system have an extremely low or null contribution to the score. Similarly, according to (3.1), master and doctor supervisions are either not considered or almost useless for the score of a researcher. In contrast, publications in periodicals ranked as $A1$ and $A2$ have, by far, the largest contributions to the score. Specifically, a publication in a journal ranked as $A1$ weights more than $1/3$ of the threshold θ while a publication in a journal ranked as $A2$ weights approximately $1/5$ of the threshold. As a consequence, if the score based model given by (1.3) is applied in a strict manner, a researcher is qualified to become a PQ2 if he has, for example, at least three publications in journals ranked as $A1$ in the last five years. We argue below, however, that some level of tolerance have to be taken into account in the decision criteria based on the score of a researcher.

Now, let us turn our attention to our curricula database. Figure 1 shows the distribution of the scores of both PQ2 and non-PQ researchers using the QUALIS ranking. Moreover, Table 1 provides some descriptive statistics computed over the score of either PQ2 fellows and non-PQ researchers. Note that the median of the scores of non-PQ researchers and PQ-fellows are,

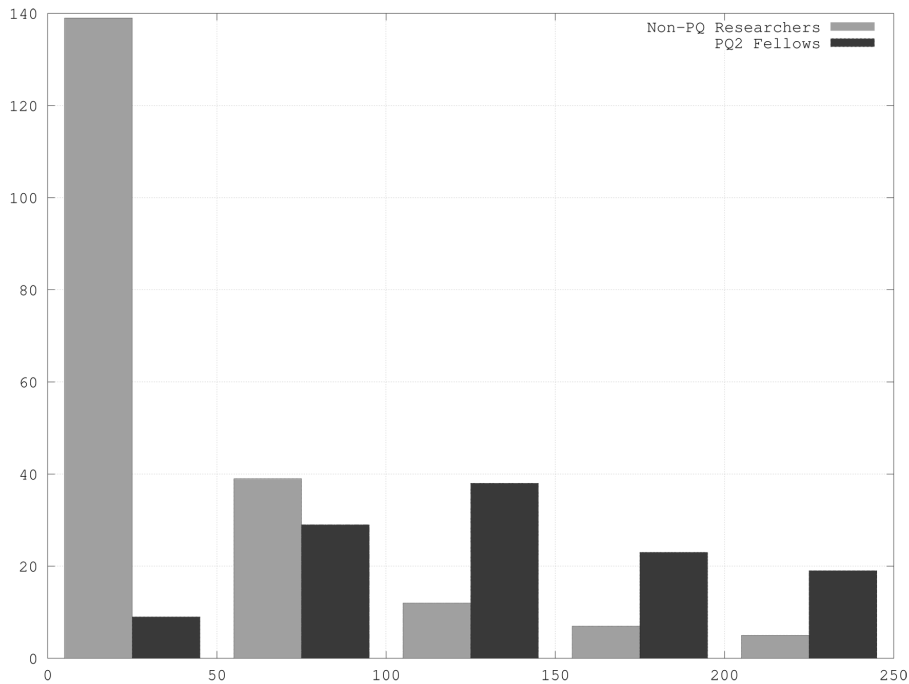


Figure 1: Distribution of the scores of PQ2 fellows (dark-gray) and non-PQ researchers (gray).

Table 1: Statistics of the score of PQ-fellows and non-PQ researchers.

	minimum	first quartile	median	third quartile	maximum
Non-PQ researchers	0.33	9.10	34.87	63.64	313.64
PQ-fellows	9.10	92.45	121.98	167.96	539.60

respectively, 34.87 and 121.98. Hence, more than half of the PQ fellows from our database have a score greater than $\theta = 100$. Also, observe that the score of the majority of non-PQ researchers is below 100. In particular, the score of 139 non-PQ researchers, which represents 68.81% of the class, is less than or equal to 50.

The objective criteria given by (1.3) with the weights and threshold obtained from the QUALIS system resulted in 62 misclassified curricula, which corresponds to an error of 19.38%. Such larger error can be partially justified by the following remarks:

- We have not considered the merit of the research project. Also, not all non-PQ fellows applied for a PQ fellowship in the last two years.
- The administrative council may take into account subjective or hard to measure information such as the potential for leadership and the regularity of the scientific production.

- We have collected the Lattes curricula in September 2014 while they have been analyzed by the administrative council of CNPq approximately one or two years before. Some curricula certainly have been modified during this period of time.
- The QUALIS system is designed to evaluate graduate programs. Hence, it may be inappropriate to appraise the scientific production of a researcher. For instance, a publication in an influential journal, such as *Nature* or *Science* may be considered by the administrative council. However, since we computed the score of a researcher using the QUALIS system, it is possible that such influential journal have not been classified in the field of mathematics and statistics yet.

The comments in the previous paragraph strongly suggest us to consider a soft version of (1.3) with the weight vector given by (3.1). By allowing a tolerance of 50 points in the criteria given by (1.3), only 21 curricula are misclassified, which corresponds to an error of 6.56%. Specifically, 9 PQ fellows would not be able to renew their fellowship because their score is below 50. The five least scores of PQ fellows are shown in Table 2 together with their number of publications and supervisions in the last five years. Similarly, observe that only 12 non-PQ researchers have score greater than or equal to 150, which represents 5.94% of the class. The score as well as the production of the five non-PQ researchers with largest scores are shown in Table 3. Finally, we would like to remark that we have not pursued an explanation for these misclassified curricula – we only confirmed that the python software have counted correctly their scientific production as well as the formation of human resources.

Table 2: Score and production of the five PQ fellows with smallest scores.

Score	A1	A2	B1	B2	B3	B4	B5	C	N	M	D
9.10	0	0	0	2	0	0	0	0	2	1	0
13.64	0	0	4	0	0	0	0	0	1	0	0
17.01	0	0	5	0	0	1	1	0	1	1	0
18.19	0	0	3	3	0	0	0	0	0	2	0
21.29	0	0	6	0	0	0	0	0	0	5	2

Table 3: Score and production of the five non-PQ researchers with largest scores.

Score	A1	A2	B1	B2	B3	B4	B5	C	N	M	D
313.64	7	2	0	0	0	0	0	0	1	3	1
257.55	6	1	0	2	0	0	0	0	1	0	0
257.55	6	1	0	2	0	0	0	0	1	0	0
245.69	1	7	5	8	0	0	1	0	2	2	4
228.87	2	5	9	1	0	0	1	0	6	6	0

4 CONCLUSION

In this paper, we modeled the decision criteria for receiving a PQ2 fellowship from CNPq in the field of mathematics. The model is based on a weighted sum of the number of publications, classified according to the QUALIS systems and the supervisory experience in graduate programs. The weights were obtained by solving a bilevel optimization problem subject to the information available at the Lattes curricula of both PQ2 fellows and non-PQ researchers. By allowing a certain tolerance, the model yielded an error of 6.56%. Briefly, the criteria based on the weights given by (3.1) can be translated into the following rule, which corresponds to a soft version of (1.3):

$$\left\{ \begin{array}{l} \text{If } 150 < S(\mathbf{x}) \text{ then the applicant } \mathbf{x} \text{ is worth to receive a PQ2 fellowship,} \\ \text{If } 50 \leq S(\mathbf{x}) \leq 150 \text{ then it is possible that the applicant receives a PQ2 fellowship,} \\ \text{If } S(\mathbf{x}) < 50 \text{ then the applicant } \mathbf{x} \text{ is not worth to receive a PQ2 fellowship.} \end{array} \right. \quad (4.1)$$

For future research, we suggest to investigate the temporal variations of the weights used to compute the score of a PQ2 applicant. Also, we encourage researchers of other disciplines to perform a similar study. In fact, since the threshold for receiving a PQ2 fellowship can be fixed at 100, the score number can be adopted in many situations to compare the productivity of researchers of different areas of knowledge.

ACKNOWLEDGMENT

We would like to thank Felipe Kenji Nakano for collecting the Lattes curricula. This work was supported in part by CNPq and FAPESP under grants nos. 304240/2011-7 and 2013/12310-4, respectively.

A AN ALTERNATIVE TO THE BILEVEL OPTIMIZATION PROBLEM

Alternatively, the bilevel optimization problem given by (2.7) and (2.8) can be replaced by the following quadratic programming problem whose objective is to optimize the margin of separation as well as the classification error:

$$\left\{ \begin{array}{l} \underset{\mathbf{w}, \mathbf{s}}{\text{minimize}} \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{\xi=1}^{320} s_{\xi} \\ \text{subject to} \quad d_{\xi} (\mathbf{w}^T \mathbf{x}_{\xi} - b) \geq 1 - s_{\xi}, \quad \forall \xi = 1, \dots, 320, \\ \quad \quad \quad w_1 \geq w_2 \geq w_3 \geq w_4 \geq w_5 \geq w_6 \geq w_7 \geq w_8, w_{11} \geq w_{10}, \\ \quad \quad \quad \mathbf{w} \in \mathbb{R}_+^{11}, b \in \mathbb{R}_+, \text{ and } \mathbf{s} \in \mathbb{R}_+^{320}, \end{array} \right. \quad (A.1)$$

where $\mathbf{s} = [s_1, s_2, \dots, s_{320}]^T$ is the vector whose components are the slack variables and $C > 0$ determines the trade-off between the margin of separation and the classification error. For instance, more emphasis is given to the classification error if C is assigned a large value. In particular, the quadratic programming problem (A.1) converged to a solution similar to \mathbf{w}^* and b^* given by (2.9) for $C \geq 1$.

RESUMO. A bolsa de produtividade (PQ) concedida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), além do aporte financeiro, fornece uma posição de destaque para pesquisadores brasileiros de todas as áreas do conhecimento. Consequentemente, tanto o perfil como os critérios para receber uma bolsa PQ são de interesse de toda comunidade científica brasileira. Neste artigo, modelamos o critério de decisão como uma soma ponderada da produção científica e das orientações de um candidato à bolsa PQ. A produção científica é determinada pelo número de publicações agrupadas de acordo com o sistema QUALIS fornecido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). O currículo Lattes de bolsistas PQ na área de matemática, juntamente com currículos de muitos pesquisadores não bolsistas de instituições semelhantes, foram usados para estimar um critério para receber bolsa PQ na categoria 2. Permitindo uma certa tolerância, o modelo reproduziu o critério adotado com limites aceitáveis sobre um base de dados com 320 currículos.

Palavras-chave: produção científica, bolsa de produtividade, matemática, otimização.

REFERENCES

- [1] D. Arruda, F. Bezerra, V.A. Neris, P.R. de Toro & J. Wainer. Brazilian Computer Science Research: Gender and Regional Distributions. *Scientometrics*, **79**(3) (2009), 651–656.
- [2] R.B. Barata & M. Goldbaum. Perfil dos pesquisadores com bolsa de produtividade em pesquisa do CNPq da área de saúde coletiva. *Cadernos de Saúde Pública*, **19**(6) (2003), 1863–1876.
- [3] S. Dempe. *Foundations of Bilevel Programming*. Kluwer Academic Publishers, (2002).
- [4] N. Garcia & L. Casado. Comunicado 003/2012 – Área de Matemática, Probabilidade e Estatística: Atualização do WebQualis. Available at: http://www.capes.gov.br/images/stories/download/avaliacao/Qualis__Matematica.pdf, 2012. Accessed on June, 2012.
- [5] S. Haykin. *Neural Networks and Learning Machines*. Prentice-Hall, Upper Saddle River, NJ, 3rd edition edition, (2009).
- [6] E.A. Oliveira, E.A. Colosimo, D.R. Martelli, I.G. Quirino, M.C.L. Oliveira, L.S. Lima, A.C.S. Silva & H. Martelli-Junior. Comparison of Brazilian researchers in clinical medicine: are criteria for ranking well-adjusted? *Scientometrics*, **90**(2) (2012), 429–443.
- [7] E.A. Oliveira, A.L.P. Ribeiro, I.G. Quirino, M.C.L. Oliveira, D.R. Martelli, L.S. Lima, E.A. Colosimo, T.J. Lopes, A.C.S. Silva & H. Martelli-Junior. Profile and Scientific Production of CNPq Researchers in Cardiology. *Arquivos Brasileiros de Cardiologia*, **97**(3) (2011), 186–193.
- [8] N.C. Santos, L.F. Cândido & C.L. Kuppens. Produtividade em Pesquisa do CNPq: Análise do Perfil dos Pesquisadores da Química. *Química Nova*, **33**(2) (2010), 489–495.
- [9] A.C. Scarpelli, F. Sardenberg, D. Goursand, S.M. Paiva & I.A. Pordeus. Academic trajectories of dental researchers receiving CNPq's productivity grants. *Brazilian Dental Journal*, **9**(3) (2008), 252–256.
- [10] M.A. Silveira e Pereira Neves, R.C. Antunes, E.S.R. Júnior & C.Q. Gorgati. Evolução das bolsas de produtividade em pesquisa e dos editais universais do CNPq no Programa Básico de Zootecnia: 2002 a 2006. *Revista Brasileira de Zootecnia*, **36**(0) (2007), 369–376.

- [11] G.A. Toffoli & S.M.S.P. Ferreira. Mapeamento da produção científica de pesquisadores brasileiros de ciências da comunicação: período de 2000 a 2009. *Psicologia USP*, **22**(2) (2011), 399–422.
- [12] V.N. Vapnik. *Statistical Learning Theory*. John Wiley and Sons, New York, NY, USA, (1998).
- [13] V.N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, 2 edition, (1999).