Journal of Uncertain Systems
Vol.4, No.4, pp.291-295, 2010
Online at: www.jus.org.uk

Towards Optimal Allocation of Points to Different Assignments and Tests

Suparat Niwitpong[1], Monchaya Chiangpradit[1], Olga Kosheleva[2,*]

[1] *Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, 1518 Pibulsongkram Rd., Bangsue, Bangkok 10800 Thailand*

[2] *Department of Teacher Education, University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA*

Received 21 December, 2008; Revised 30 January, 2010

Abstract

How many points should we allocate to different assignments and tests? to different problems on a test? Usually, professors use subjective judgment to allocate points. In this paper, we provide an objective procedure for allocating points.

Keywords: grading, optimization

1 Formulation of the Problem

Points need to be allocated. The overall grade for a class is formed by adding the grades for different assignments and tests. In the syllabus, it is usually described that, e.g., the first test is worth 10 points, the second test is worth 25 points, etc., with the total of 100 points.

Example. For a better understanding, let us start with a simple example. Suppose that we allocate:

- 10 points to the first test,
- 40 points to the second test, and
- 50 points to the final exam.

Suppose that a student s got:

- $g_{s1} = 80$ points out of 100 on the first test,
- $g_{s2} = 75$ points out of 100 on the second test, and
- $g_{s3} = 90$ points out of 100 on the final exam.

In this case, the student's overall grade g_s for this class is

$$g_s = p_1 \cdot g_{s1} + p_2 \cdot g_{s2} + p_3 \cdot g_{s3} = 0.1 \cdot 80 + 0.4 \cdot 75 + 0.5 \cdot 90 = 8.0 + 30.0 + 45.0 = 83, \tag{1}$$

where

$$p_1 = \frac{10}{100} = 0.1, \quad p_2 = \frac{40}{100} = 0.4, \quad p_3 = \frac{50}{100} = 0.5. \tag{2}$$

Points need to be allocated (cont-d). Similarly, the grade for a test is formed by adding the grades earned on each of the problems. In the text of the test, it is usually described how many points each problem is worth.

*Corresponding author. Email: olgak@utep.edu (O. Kosheleva).

Allocating points is important. An appropriate allocation of points is very important. For example, if we allocate almost all the points to the final exam, some students will see no reason to study hard during the semester. They will try to cram the material during the finals, and as a result, even if they pass the finals, their knowledge of the material will not be as good as the knowledge of the students who studied diligently during the semester.

On the other hand, if we allocate too few points to the final exam, then some students will have no incentive to review the class material for the final exam. As a result, their knowledge of the material that was studied in the beginning of the semester will be not as good as the knowledge of those students who did review this material at the end of the class.

How points are allocated now. At present, the points for different assignments, tests, and problems are allocated based on the teacher's subjective experience. As a result, there is a high variety of point allocation.

Formulation of the problem. Since it is very important to properly allocate points for different assignments, tests, and problems, it is desirable to find a more objective ways for such allocation.

Such an objective way is presented in this paper. This paper capitalizes on the preliminary ideas described in [1].

2 A New Approach

Our main idea. Most classes are prerequisites for other classes (or for some qualification exams). For example, Pre-calculus is a prerequisite for Calculus I, Calculus I is a prerequisite for Calculus II, etc. For such classes, it is desirable to allocate the points in such a way that a success in this class will be a good indication of the success in the next class.

If we grade too easily, students may be happier with their good grades, but some of them will pass the class without acquiring the knowledge needed for the next class – and so they may fail this next class. On the other hand, if we grade too harsh, we unnecessarily fail many students who may have not learned all the details but whole knowledge is actually good enough to successfully pass the next class.

From this viewpoint, the best way to allocate points is to select the allocations for which the resulting grade is the best predictor for the grade in the next class.

Available data. To find the best allocation of points, we must use the grades that the students got for different assignments and tests, and the grades they got in the next class.

Let us denote the total number of assignments and tests by T, and let us denote the total number of students who took this class in the past by S. Let us denote the grade of student s ($1 \le s \le S$) on assignment or test t ($1 \le t \le T$) by g_{st}. The grade of student s at the next class (for which this class is a prerequisite) will be denoted by n_s.

From the idea to the algorithm. We want to predict the value n_s based on the grades $g_{s1}, \ldots, g_{sT}$. For such a prediction, it is natural to start with a linear regression

$$n_s \approx a_0 + a_1 \cdot g_{s1} + \ldots + a_T \cdot g_{sT}. \tag{3}$$

The coefficients a_t can be found from the Least Squares method (see, e.g., [2]), by minimizing the sum

$$\sum_{s=1}^{S} (a_0 + a_1 \cdot g_{s1} + \ldots + a_T \cdot g_{sT} - n_s)^2. \tag{4}$$

The standard Least Squares formulas lead to the following system of linear equations for determining the coefficients a_t:

$$a_0 + a_1 \cdot \overline{g_1} + \ldots + a_T \cdot \overline{g_T} = \overline{n}; \tag{5}$$

$$a_t + a_1 \cdot \overline{g_1 \cdot g_t} + \ldots + a_T \cdot \overline{g_T \cdot g_t} = \overline{n \cdot g_t} \quad 1 \le t \le T, \tag{6}$$

where

$$\overline{g_t} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} g_{st}, \quad \overline{n} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} n_s, \quad \overline{g_t \cdot g_{t'}} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} g_{st} \cdot g_{st'}. \tag{7}$$

The resulting coefficients are not exactly points, since the values $a_1, \ldots, a_T$ do not necessarily add up to 1. To get the desired values of the points $p_1, \ldots, p_T$, we therefore need to normalize these values and take

$$p_t = \frac{a_t}{a}, \tag{8}$$

where

$$a \stackrel{\text{def}}{=} a_1 + \ldots + a_T. \tag{9}$$

For these values, we have $a_t = p_t \cdot a$ and thus, the above linear regression formula takes the form

$$n_s \approx a_0 + a_1 \cdot g_{s1} + \ldots + a_T \cdot g_{sT} = a_0 + a \cdot g_s, \tag{10}$$

where

$$g_s \stackrel{\text{def}}{=} p_1 \cdot g_{s1} + \ldots + p_T \cdot g_{sT}. \tag{11}$$

Thus, we arrive at the following algorithm.

Allocating points: main algorithm. Based on the grades $g_{s1}, \ldots, g_{sT}$, and n_s, we first compute the averages

$$\overline{g_t} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} g_{st}, \quad \overline{n} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} n_s, \quad \overline{g_t \cdot g_{t'}} \stackrel{\text{def}}{=} \frac{1}{S}\sum_{s=1}^{S} g_{st} \cdot g_{st'}. \tag{12}$$

Then, we solve the following system of $T+1$ linear equations to find $T+1$ unknowns $a_0, a_1 \ldots, a_T$:

$$a_0 + a_1 \cdot \overline{g_1} + \ldots + a_T \cdot \overline{g_T} = \overline{n}; \tag{13}$$

$$a_t + a_1 \cdot \overline{g_1 \cdot g_t} + \ldots + a_T \cdot \overline{g_T \cdot g_t} = \overline{n \cdot g_t} \quad 1 \le t \le T. \tag{14}$$

After that, we compute the sum

$$a \stackrel{\text{def}}{=} a_1 + \ldots + a_T \tag{15}$$

and we compute the desired point values p_t as

$$p_t = \frac{a_t}{A}. \tag{16}$$

For each student s, the resulting class grade

$$g_s \stackrel{\text{def}}{=} p_1 \cdot g_{s1} + \ldots + p_T \cdot g_{sT} \tag{17}$$

can be used to predict the grade n_s in the next class as

$$n_s \approx a_0 + a \cdot g_s. \tag{18}$$

Discussion. One might expect that we would aim at finding the points for which $n_s \approx g_s$. However, it is reasonable to expect some decrease in knowledge between the classes, so $n_s \approx a_0 + a \cdot g_s$, with $a < 1$, is a more reasonable idea.

For example, in the University of Texas at El Paso's Computer Science (CS) undergraduate program, the passing grade for some classes is D (corresponding, crudely speaking, to 60 points out of 100). However, for the classes that serve as prerequisites to other classes, the passing grade is C (70 points out of 100), because it is understood that usually, student forget. So, to make sure that they have at least a D level knowledge of the material by the time they reach the next class, we must make sure that they have a C level by the time they finish the previous class.

3 Additional Complications

Two iterations may be needed. In the previous text, we assumed that the grades for different assignments reflect the students' level of knowledge for different parts of the material. This is usually true when there are enough points allocated to this assignment. However, if too few points are allocated to an assignment, students may not spend as much time on it and thus, the corresponding grades may be low. If based on the above procedure, we allocate more points to this assignment, students may take it more seriously and their grades for this assignment may improve. This change in grades may require a re-allocation of points.

So, ideally, after we determine the initial grade allocations, we should then collect new grades data. After we have collected enough grade data, we should then repeat the same procedure – to see if re-allocation of points is necessary.

From the main algorithm to algorithms that cover more complex situations. In the derivation of the above algorithm, we assumed that we have a class with a well-defined next class, and that for this next class, the current class is the only prerequisite. We also assumed that linear regression is adequate.

Let us discuss what to do in more complex situations.

Case of multiple prerequisites. In some situations, there are several prerequisites for a class. For example, in the above-mentioned CS program, there are two prerequisites for the Data Structures class: Discrete Mathematics (DM) and Elementary Algorithms and Data Structures (CS2).

In such cases, we need to determine points for both prerequisite classes. Here, in addition to the grades $g_{s1}, \ldots, g_{sT}$ for the first class, we also have grades $g'_{s1}, \ldots, g'_{sT'}$ for the second class. We want to find the coefficients $a_0, a_1, \ldots, a_T, a'_1, \ldots, a'_{T'}$ for which

$$n_s \approx a_0 + a_1 \cdot g_{s1} + \ldots + a_T \cdot g_{sT} + a'_1 \cdot g'_{s1} + \ldots + a'_{T'} \cdot g_{sT'}. \tag{19}$$

The coefficients a_t and $a'_{t'}$ can be found from the Least Squares method, by minimizing the sum

$$\sum_{s=1}^{S} (a_0 + a_1 \cdot g_{s1} + \ldots + a_T \cdot g_{sT} + a'_1 \cdot g'_{s1} + \ldots + a'_{T'} \cdot g'_{sT'} - n_s)^2. \tag{20}$$

The standard Least Squares formulas lead to the following system of linear equations for determining the coefficients a_t and $a'_{t'}$:

$$a_0 + a_1 \cdot \overline{g_1} + \ldots + a_T \cdot \overline{g_T} + a'_1 \cdot \overline{g'_1} + \ldots + a_T \cdot \overline{g'_{T'}} = \overline{n}; \tag{21}$$

$$a_t + a_1 \cdot \overline{g_1 \cdot g_t} + \ldots + a_T \cdot \overline{g_T \cdot g_t} + a'_1 \cdot \overline{g'_1 \cdot g_t} + \ldots + a'_{T'} \cdot \overline{g'_{T'} \cdot g_t} = \overline{n \cdot g_t}; \tag{22}$$

$$a'_{t'} + a_1 \cdot \overline{g_1 \cdot g'_{t'}} + \ldots + a_T \cdot \overline{g_T \cdot g'_{t'}} + a'_1 \cdot \overline{g'_1 \cdot g'_{t'}} + \ldots + a'_{T'} \cdot \overline{g'_{T'} \cdot g'_{t'}} = \overline{n \cdot g'_{t'}}, \tag{23}$$

where, in addition to (7), we have

$$\overline{g'_{t'}} \stackrel{\text{def}}{=} \frac{1}{S} \sum_{s=1}^{S} g'_{st'}, \quad \overline{g_t \cdot g'_{t'}} = \overline{g'_{t'} \cdot g_t} \stackrel{\text{def}}{=} \frac{1}{S} \sum_{s=1}^{S} g_{st} \cdot g'_{st'}, \tag{24}$$

$$\overline{g'_t \cdot g'_{t'}} \stackrel{\text{def}}{=} \frac{1}{S} \sum_{s=1}^{S} g'_{st} \cdot g'_{st'}. \tag{25}$$

To get the desired values of the points $p_1, \ldots, p_T, p'_1, \ldots, p'_{T'}$, we normalize these values and take

$$p_t = \frac{a_t}{a}, \quad p'_{t'} = \frac{a'_{t'}}{a'}, \tag{26}$$

where

$$a \stackrel{\text{def}}{=} a_1 + \ldots + a_T; \quad a' \stackrel{\text{def}}{=} a'_1 + \ldots + a'_{T'}. \tag{27}$$

For these values, we have $a_t = p_t \cdot a$ and $a'_{t'} = p'_{t'} \cdot a'$ and thus, the above linear regression formula takes the form

$$n_s \approx a_0 + a \cdot g_s + a' \cdot g'_s, \tag{28}$$

where

$$g_s \stackrel{\text{def}}{=} p_1 \cdot g_{s1} + \ldots + p_T \cdot g_{sT}; \quad g'_s \stackrel{\text{def}}{=} p'_1 \cdot g'_{s1} + \ldots + p'_T \cdot g'_{sT'}. \tag{29}$$

Case of several follow-up classes. Another realistic situation is when for a given class, there are several follow-up classes. For example, in the above-mentioned CS program, the Data Structures class is a pre-requisite for several junior- and senior-level classes. In this case, for each student s, we have several grades $n_{s1}, \ldots, n_{sC}$ for the follow-up classes.

To apply the above procedure, we can then take, as n_s,

- either the *smallest* of these follow-up grades
$$n_s = \min(n_{s1}, \ldots, n_{sC}); \tag{30}$$
this will guarantee that the student is successful in *all* follow-up classes,

- or the *average* of these follow-up grades
$$n_s = \frac{n_{s1} + \ldots + n_{sC}}{C}; \tag{31}$$
this will guarantee that the student is successful *on average* in the follow-up classes.

Case of non-linear regression. If we have reasons to believe that linear regression does not adequately capture the students' success in the next classes, we may want to use quadratic regression, i.e., find the coefficients a_t and $a_{tt'}$ for which

$$n_s \approx a_0 + \sum_{t=1}^{T} a_t \cdot g_{st} + \sum_{t=1}^{T} \sum_{t'=1}^{T} a_{tt'} \cdot g_{st} \cdot g_{st'}. \tag{32}$$

The optimal values of the coefficients a_t and $a_{tt'}$ can also be obtained from the Least Squares method. In this case, the grade for the class will be obtained not as a weighted average of the grades for different assignments and tests, but rather as a non-linear combination of these grades.

This non-linearity may sound unusual, but it is actually used in grading. For example, in some CS classes, to get a C, students need to get at least a C average for the lab assignments (assignments $t = 1, \ldots, \ell$) and at least a C average for all the tests (assignments $t = \ell + 1, \ldots, T$). Thus, in effect, we have a non-linear formula for the class grade:

$$g_s = \min\left(\frac{g_{s1} + \ldots + g_{s\ell}}{\ell}, \frac{g_{s,\ell+1} + \ldots + g_{s,T}}{T - \ell}\right). \tag{33}$$

Acknowledgments

The authors are thankful to Sa-aat Niwitpong and to Vladik Kreinovich for their guidance and help.

References

[1] Kosheleva, O., and V. Kreinovich, How to assigns weights to different tests?, *Proceedings of the Sun Conference on Teaching and Learning*, El Paso, Texas, February 27, 2009.

[2] Sheskin, D.J., *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC Press, Boca Raton, Florida, 2007.