The Career Atlas: Mathematical Notation

Cao Bittencourt

[a] *B. Sc. in Economics from EPGE (FGV), RJ, Brazil.*
[b] *Statistician at Atlas Career Guide Inc., FL, USA.*

Abstract

This is a brief document to define statistical methods for data-drive career choice and development. It deals with topics such as: career matching (i.e. vocational choice); estimation of competence, or overall skill level; estimation of skill set generality; versatility; skill set profitability; employability; labor market competitiveness; labor market taxonomy; optimal human resources acquistion and allocation; and so on and so forth. Each concept shall be explained at length in separate articles.

Keywords: Career choice; Career development; Matching algorithms; Competence; Similarity.

1. Basic Definitions

1.1. Skill Sets

The i-th professional attribute, or competency, of a person k is defined as:

$$a_i^k \in [0, 100], \tag{1}$$

where the interval $[0, 100]$ determines the bounds for every competency.[1]

The skill set, or career profile, of a person k is defined as the vector of their m attributes:

$$\boldsymbol{a_k} = (a_1^k, \dots, a_m^k). \tag{2}$$

A skill set matrix, or career profile matrix, is the collection of all n skill sets in the economy:

$$\mathbf{A} = \begin{bmatrix} a_1^1 & \dots & a_m^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \dots & a_m^n \end{bmatrix}. \tag{3}$$

[1] More generally, these could be defined as a_{lb} (the lower bound) and a_{ub} (the upper bound). Here, the interval $[0, 100]$ is used because of its ease of interpretation.

1.2. Skill Set Normalization

Normalization by the scale bounds is defined by the tilde operator:

$$\tilde{a}_i^k = \frac{a_i^k - 0}{100 - 0} = \frac{a_i^k}{100} \in [0, 1]; \tag{4}$$

$$\tilde{\boldsymbol{a}}_{\boldsymbol{k}} = (\tilde{a}_1^k, \ldots, \tilde{a}_m^k); \tag{5}$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{a}_1^1 & \ldots & \tilde{a}_m^1 \\ \vdots & \ddots & \vdots \\ \tilde{a}_1^n & \ldots & \tilde{a}_m^n \end{bmatrix}. \tag{6}$$

Normalization by a skill set's highest attribute is defined by the hat operator:

$$\hat{a}_i^k = \frac{a_i^k}{\max a_j^k} \in [0, 1]; \tag{7}$$

$$\hat{\boldsymbol{a}}_{\boldsymbol{k}} = (\hat{a}_1^k, \ldots, \hat{a}_m^k); \tag{8}$$

$$\hat{\mathbf{A}} = \begin{bmatrix} \hat{a}_1^1 & \ldots & \hat{a}_m^1 \\ \vdots & \ddots & \vdots \\ \hat{a}_1^n & \ldots & \hat{a}_m^n \end{bmatrix}. \tag{9}$$

2. Basic Skill Set Functions

The generality of a skill set is the mean of its maxima-normalized attributes:

$$\gamma_k = \left(\frac{1}{m}\right) \sum_{i=1}^{m} \hat{a}_i^k. \tag{10}$$

Generalists have high γ_k scores. Specialists have low γ_k scores. The generality vector of all n skill sets in the economy is:

$$\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n). \tag{11}$$

The attribute equivalence of a particular attribute in a skill set measures the importance of that attribute relative to the skill set's highest attribute, using the skill set's generality as a midpoint and scaling parameter. The attribute equivalence of an attribute is denoted by the umlaut operator:

$$\ddot{a}_i^k = \text{aeq}(\hat{a}_i^k, \gamma_k) = \hat{a}_i^k \left[1 + \gamma_k (1 - \hat{a}_i^k) \exp\left(\frac{\hat{a}_i^k - \gamma_k}{\gamma_k - 1}\right)\right]^{-\frac{\gamma_k}{\hat{a}_i^k}}. \tag{12}$$

Attributes with high levels of attribute equivalence $\ddot{a}_i^k$ are said to be equivalent to the skill set's most importante attribute. These attributes are called *core* attributes. The attribute equivalence vector of a skill set is given by the collection of their m umlauted attributes:

$$\ddot{\boldsymbol{a}}_{\boldsymbol{k}} = (\ddot{a}_1^k, ..., \ddot{a}_m^k). \tag{13}$$

The attribute equivalence matrix is the collection of all attribute equivalence vectors in the economy:

$$\ddot{\mathbf{A}} = \begin{bmatrix} \ddot{a}_1^1 & \dots & \ddot{a}_m^1 \\ \vdots & \ddots & \vdots \\ \ddot{a}_1^n & \dots & \ddot{a}_m^n \end{bmatrix}. \tag{14}$$

The overall competence of a skill set is the mean of its scale-normalized attributes, weighted by each attribute's importance (i.e. its attribute equivalence):

$$c_k = \frac{\sum_{i=1}^{m} \ddot{a}_i^k \tilde{a}_i^k}{\sum_{i=1}^{m} \ddot{a}_i^k}. \tag{15}$$

The competence vector of all n skill sets in the economy is:

$$\boldsymbol{c} = (c_1, \dots, c_n). \tag{16}$$