# 36. Multidimensional Quantification. II

By Chikio HAYASHI
Institute of Statistical Mathematics, Tokyo
(Comm. by Z. SUETUNA, M.J.A., March 12, 1954)

The present paper is a continuation of [3] and deeply related to [1], [2]. The construction of problems, the meanings of symbols, terms and definitions etc. are the same as in [3]. For simplicity we call the method of [3] "case 1".

§ 1. Case 2, where elements are classified into $S$ strata by an outside criterion which is not unidimensional. Each element has response patterns in $R$ items and the label of the stratum to which it belongs. This label is called an outside variable. We use the same symbols as in [3]. In this case we consider the $R$-dimensional Euclidean space, each dimension of which corresponds to each item. Now we set $n$ orthogonal axes corresponding to the items. Each element will be represented as a point in this space if sub-categories in items are to be quantified. Now we should like to quantify the sub-categories $C_{tm}$ so as to maximize the effect of stratification. Here it is not at all effective to use the method of the case 1. As the total variance $\sigma^2$, we take the generalized variance which is considered to be proportionate to square of the volume of the so-called ellipsoid of concentration, where $\sigma^2=|\rho_{jl}\sigma_j\sigma_l|$, $\rho_{jl}\sigma_j\sigma_l$ is covariance between the $j$-th and the $l$-th item (dimension) when the sub-categories are quantified and $|\cdots|$ expresses a determinant, the element of which is $\rho_{jl}\sigma_j\sigma_l$ $j, l=1, 2,\ldots, R$. As the within variance, we take $\sigma_t^2=|\rho_{jl}(t)\sigma_j(t)\sigma_l(t)|$, where $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)$ is covariance between the $j$-th and the $l$-th item (dimension) in the $t$-th stratum, which is deeply related, in the above sense, to the ellipsoid of concentration in the $t$-th stratum.

Thus we take $\mu=1-(\sum_{t=1}^{S} p_t\sigma_t^2/\sigma^2)$ as a measure of efficiency of stratification, i.e. a measure of discriminative power of items which will be an index related to the efficiency of classification (success rate of prediction) by quantified behaviour patterns, where $p_t$ is a weight assigned to the $t$-th stratum, $\sum_{t=1}^{S} p_t=1$; especially in this case, we take $p_t=\frac{n_t}{n}$, $n_t$ is the size of the $t$-th stratum, $\sum_{t=1}^{S} n_t=n$. $\sum_{t=1}^{S} p_t\sigma_t^2$ is considered to be a sort of so-called within variance in the whole. If $\sigma_t^2=0$, $\mu=1$, if $\sigma_t^2=\sigma^2$, $\mu=0$. Besides this, we can take several indices as a measure. One example of these will be shown later on.

We require the value $x_{lm}$ given to $C_{lm}$ to maximize $\mu$:

$$\frac{\partial\mu}{\partial x_{uv}}=0,\ u=1,2,\dots,R;\ v=1,2,\dots,K_u.$$

So we obtain

$$\sum_{t=1}^{S}p_t\frac{\partial\sigma_t^2}{\partial x_{uv}}=(1-\mu)\frac{\partial\sigma^2}{\partial x_{uv}}. \tag{1}$$

It is our problem to solve these equations and obtain the largest maximum value of $\mu(\neq 1)$ (which is the largest value of $\mu$) and the corresponding vectors. Let $\bar{x}_j$ be the total mean of the $j$-th item, $\bar{x}_j=\frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K_j}x_{jk}\delta_i(jk)=\frac{1}{n}\sum_{k}^{K_j}x_{jk}n_{jk}$, $\sigma_j^2$ be the total variance of the $j$-th item, $\sigma_j^2=\frac{1}{n}\sum_{k=1}^{K_j}x^2_{jk}n_{jk}-\bar{x}_j^2$, $\bar{x}_j(t)$ be the mean of the $j$-th item in the $t$-th stratum, $\bar{x}_j(t)=\frac{1}{n_t}\sum_{k=1}^{K_j}x_{jk}g^t(jk)$, $\sigma_j(t)^2$ be the variance of the $j$-th item in the $t$-th stratum $\sigma_j(t)^2=\frac{1}{n_t}\sum_{k=1}^{K_j}x_{jk}^2\,g^t(jk)-\bar{x}_j(t)^2$. Without loss of generality, $\bar{x}_j=0$, $j=1,2,\dots,R$.

Thus $\sigma_j=\frac{1}{n}\sum_{k=1}^{K_j}x^2_{jk}n_{jk}$. Let $\rho_{jl}\sigma_j\sigma_l$ be the total covariance between $j$ and $l$, $\rho_{jl}\sigma_j\sigma_l=\frac{1}{n}\sum_{k}^{K_j}\sum_{m}^{K_l}x_{jk}x_{lm}f_{jk}(lm)$; $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)$ be the covariance in the $t$-th stratum, $\rho_{jl}(t)\sigma_j(t)\sigma_l(t)=\frac{1}{n_t}\sum_{k}^{K_j}\sum_{m}^{K_l}x_{jk}x_{lm}f^t_{jk}(lm)-\bar{x}_j(t)\,\bar{x}_l(t)$, where

$$f^t_{jk}(lm)=\sum_{i(t)=1}^{n_t}\delta_{i(t)}(jk)\,\delta_{i(t)}(lm),\quad \sum_{t=1}^{S}f^t_{jk}(lm)=f_{jk}(lm).$$

Thus

$$\sigma^2=\left|\frac{1}{n}\sum_{k}^{K_j}\sum_{m}^{K_l}x_{jk}x_{lm}f_{jk}(lm)\right| \tag{2}$$

$$\sigma_t^2=\left|\frac{1}{n_t}\sum_{k}^{K_j}\sum_{m}^{K_l}x_{jk}x_{lm}f^t_{jk}(lm)-\bar{x}_j(t)\bar{x}_l(t)\right|. \tag{3}$$

From these (1), (2), (3), we can obtain the numerical values $x_{lm}$ ($l=1,2,\dots,R$, $m=1,2,\dots,K_l$) which maximize $\mu$.

Special case: Let $R=3$.

$$\sigma^2=\begin{vmatrix}\sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3\\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3\\ \rho_{13}\sigma_{13}\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2\end{vmatrix};$$

$$\frac{\partial\sigma^2}{\partial x_{\alpha u}}=2(x_{\alpha u}\,p_{\alpha u}\,\tau^2_{\beta\gamma}-\sum_{k=1}^{K_\alpha}x_{\alpha k}\gamma_{\alpha,ku}),$$

$$\alpha=1,2,3,\quad \alpha\neq\beta,\quad \beta\neq\gamma,\quad \alpha\neq\gamma;\quad \beta,\gamma=1,2,3,$$

where

$$\tau^2_{\beta\gamma}=\sigma^2_\beta\sigma^2_\gamma-\rho^2_{\beta\gamma}\sigma^2_\beta\sigma^2_\gamma,\qquad (\beta,\gamma=1,2,3),$$

$$\gamma_{\alpha,ku}=S_k(\alpha\gamma)S_u(\alpha\gamma)\sigma_\beta^2+S_k(\alpha\beta)S_u(\alpha\beta)\sigma_\gamma^2$$
$$-\{S_u(\alpha\beta)S_k(\alpha\gamma)+S_u(\alpha\gamma)S_k(\alpha\beta)\}\rho_{\beta\gamma}\sigma_\beta\sigma_\gamma,$$
$$\alpha\neq\beta,\quad \alpha\neq\gamma,\quad \beta\neq\gamma,\quad \alpha=1,2,3,$$
$$S_k(\delta\varepsilon)=\frac{1}{n}\sum_m^{K_\varepsilon} x_{\varepsilon m}f_{\delta k}(\varepsilon m),\qquad \delta,\varepsilon=1,2,3,$$
$$p_{\alpha u}=\frac{n_{\alpha u}}{n}.$$

We have

$$\frac{\partial\sigma_t^2}{\partial x_{\alpha u}}=2\left\{\tau_{\beta\gamma}(t)^2(x_{\alpha u}p_{\alpha u}^t-p_{\alpha u}^t\sum_k^{K_\alpha}x_{\alpha k}p_{\alpha k}^t)-\sum_k^{K_\alpha}x_{\alpha k}\gamma_{\alpha,ku}^t\right\},$$
$$\tau_{\beta\gamma}(t)^2=\sigma_\beta(t)^2\sigma_\gamma(t)^2-\rho_{\beta\gamma}(t)^2\sigma_\beta(t)^2\sigma_\gamma(t)^2,$$
$$\gamma_{\alpha,ku}^t=S_k^t(\alpha\gamma)S_u^t(\alpha\gamma)\sigma_\beta(t)^2+S_k^t(\alpha\beta)S_u^t(\alpha\beta)\sigma_\gamma(t)^2$$
$$-\{S_u^t(\alpha\beta)S_k^t(\alpha\gamma)+S_u^t(\alpha\gamma)S_k^t(\alpha\beta)\}\rho_{\beta\gamma}(t)\sigma_\beta(t)\sigma_\gamma(t),$$
$$S_k^t(\delta\varepsilon)=\frac{1}{n_t}\sum_{m=1}^{K_\varepsilon}x_{\varepsilon m}f_{\delta k}^t(\varepsilon m)-\frac{g^t(\delta k)}{n_t}\bar{x}_\varepsilon(t),$$
$$p_{\alpha k}^t=\frac{g^t(\alpha k)}{n_t},\quad \alpha,\beta,\gamma=1,2,3,\quad \alpha\neq\beta,\quad \alpha\neq\gamma,\quad \beta\neq\gamma,\quad \delta,\varepsilon=1,2,3.$$

So in the long run, we obtain

$$\sum_{k=1}^{K_1}\Lambda_{1,ku}\,x_{1k}=\mu\sum_{k=1}^{K_1}\Phi_{1,ku}\,x_{1k},\qquad (u=1,2,\dots,K_1), \tag{4}$$
$$\sum_{m=1}^{K_2}\Lambda_{2,mv}\,x_{2m}=\mu\sum_{m=1}^{K_2}\Phi_{2,mv}\,x_{2m},\qquad (v=1,2,\dots,K_2), \tag{5}$$
$$\sum_{l=1}^{K_3}\Lambda_{3,lw}\,x_{3l}=\mu\sum_{l=1}^{K_3}\Phi_{3,lw}\,x_{3l},\qquad (w=1,2,\dots,K_3), \tag{6}$$

where

$$\Lambda_{\alpha,ku}=\bar{\gamma}_{\alpha,ku}+\varphi_{\alpha,ku}-\gamma_{\alpha,ku}+\delta_{ku}(p_{\alpha u}\tau_{\beta\gamma}^2-\psi_{\alpha u}),$$
$$\Phi_{\alpha,ku}=\delta_{ku}\,p_{\alpha u}\,\tau_{\beta\gamma}^2-\gamma_{\alpha,ku},$$
$$\varphi_{\alpha,ku}=\sum_{t=1}^{S}p_t\,p_{\alpha u}^t\,p_{\alpha k}^t\,\tau_{\beta\gamma}(t)^2,$$
$$\bar{\gamma}_{\alpha,ku}=\sum_{t=1}^{S}p_t\gamma_{\alpha,ku}^t,$$
$$\psi_{\alpha u}=\sum_{t=1}^{S}p_t\,p_{\alpha u}^t\,\tau_{\beta\gamma}(t)^2,\quad \alpha=1,2,3,\quad \alpha\neq\beta,\quad \beta\neq\gamma,\quad \alpha\neq\gamma,$$

$\delta_{ku}$ is Kronecker's symbol ($\delta_{ku}=1$, $k=u$; $=0$, $k\neq u$).

$\Lambda_{\alpha,ku}$ is symmetric and the quadratic form from symmetric matrix $(\Phi_{\alpha,ku})$ is positive definite (equal to variance). So it is shown that the required $\mu$ exists. Numerical values $x_{lm}$ (in appropriate units) obtained in the valid sense by this method have the same meanings and function as in the case 1. Furthermore we can give the distances between strata explicitly in those senses which will allow useful interpretation. The above equations can be solved by successive approximation "step by step in turn" methods. Thus we can generally obtain our required results. As another measure, we can also take $\lambda=1-\frac{\sigma^2_w}{\sigma^2}$, where $\sigma^2_w=|\sum p_t\rho_{jl}(t)\sigma_j(t)\sigma_l(t)|$, $\sigma^2$ being generalized variance, which is suggested by Mr. Akaike. The method to solve is the same as above. Generally it turns out to be better that we adopt, as a first approximation, the values of $x_{lm}$ obtained by solving the following equations, for example, $\sum_k^{K_\alpha}\sum_{t=1}^{S}\frac{g^t(\alpha k)g^t(\alpha u)}{n_t}x_{\alpha k}=\eta^2 n_{\alpha u}x_{\alpha u}$, $(\alpha=1, 2, 3)$, which mean the quantification in the case where each dimension is treated separately; see case 1. The applications of this method to analyse social phenomena will be shown in [4], [6]. The method of prediction is similar to the case 1. The essential point of this method is not to use the sum of the response patterns (additivity). This is also applicable to the case where the method of case 1 is not effective even when outside criterion is unidimensional or $S=2$.

§2. Case 3, i.e. compound case. We divide $R$ items into $T$ sub-groups. Suppose that, in each sub-group, the items belonging to it fulfil the property of additivity in valid sense (see case 1), i.e. they are on unidimensional scale. In this case we regard each element as a point in $T$ dimensional Euclidean space if sub-categories of items are quantified and synthesised in sub-groups. In each dimension, we quantify sub-categories of items under the assumption of linear form by the same idea as in case 1, but the numerical values must be required in correlation with the whole. Elements are classified into $S$ strata by an outside criterion which is not unidimensional.

Let $R1$, $R2,\ldots,$ $R_T$ be the number of items in sub-groups respectively, $\sum_{r=1}^{T}R_r=R$.

Let $\sum_{u_r}^{R_r}\sum_{v_r}x_{u_rv_r}\delta_i(u_rv_r)=\alpha_i(r)$ be the values of the $i$-element in the $r$-th sub-group, where $u_r$, $v_r\ldots$ express the items and sub-categories belonging to the $r$-th sub-group and the symbols have the same meaning as in cases 1, 2.

Thus $\bar{x}_r=\frac{1}{n}\sum_i^n\alpha_i(r)=\frac{1}{n}\sum_{u_r}^{R_r}\sum_{v_r}n_{u_rv_r}x_{u_rv_r}$,

$$\sigma_r^2=\frac{1}{n}\sum_{l_r}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{v_r}x_{u_rv_r}x_{l_rm_r}f_{u_rv_r}(l_rm_r)-\bar{x}_r^2 \quad \text{in the total,}$$

$$\bar{x}_r(t)=\frac{1}{n_t}\sum_{u_r}^{R_r}\sum_{v_r}g^t(u_rv_r)x_{u_rv_r},$$

$$\sigma_r(t)^2=\frac{1}{n_t}\sum_{l_r}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{v_r}x_{u_rv_r}x_{l_rm_r}f^t{}_{u_rv_r}(l_rm_r)-\bar{x}_r(t)^2$$

in the $t$-th stratum. Without loss of generality $\bar{x}_r=0$, $r=1, 2, \ldots, T$. Thus $\sigma_r^2=\frac{1}{n}\sum_{l_r}^{R_r}\sum_{m_r}\sum_{u_r}^{R_r}\sum_{v_r} x_{u_r v_r} x_{l_r m_r} f_{u_r v_r}(l_r m_r)$. Let $\rho_{rs}\sigma_r\sigma_s$ be the total covariance between the $r$-th dimension and the $s$-th dimension,

$$\rho_{rs}\sigma_r\sigma_s=\frac{1}{n}\sum_{i}^{n}\alpha_i(r)\alpha_i(s)$$
$$=\frac{1}{n}\sum_{l_s}^{R_s}\sum_{m_s}\sum_{u_r}^{R_r}\sum_{v_r} x_{u_r v_r} x_{l_s m_s} f_{u_r v_r}(l_s m_s),$$

$\rho_{rs}(t)\sigma_r(t)\sigma_s(t)$ be the covariance in the $t$-th stratum,

$$\rho_{rs}(t)\sigma_r(t)\sigma_s(t)=\frac{1}{n_t}\sum_{l_s}^{R_s}\sum_{m_s}\sum_{u_r}^{R_r}\sum_{v_r} x_{u_r v_r} x_{l_s m_s} f^t{}_{u_r v_r}(l_s m_s)-\bar{x}_r(t)\bar{x}_s(t).$$

Thus

$$\sigma^2=|\rho_{rs}\sigma_r\sigma_s|=\left|\frac{1}{n}\sum\sum\sum\sum x_{u_r v_r} x_{l_s m_s} f_{u_r v_r}(l_s m_s)\right|,$$

$$\sigma_t^2=|\rho_{rs}(t)\sigma_r(t)\sigma_s(t)|=\left|\frac{1}{n_t}\sum\sum\sum\sum x_{u_r v_r} x_{l_s m_s} f^t{}_{u_r v_r}(l_s m_s)-\bar{x}_r(t)\bar{x}_s(t)\right|.$$

From these, we can require the numerical values $x_{j_r k_r}$ $r=1, 2, \ldots, T$ $j_r=1, 2, \ldots, R_r$, $k_r=1, 2, \ldots, K_{j_r}$ to maximize $\mu$ or $\lambda$, which have the same meaning and content as in cases 1, 2. In special cases, we can explicitly describe the equations which have the same forms as in case 2 combined with those of case 1, so that in these cases we can obtain $x_{j_r k_r}$ by the same operations using successive approximation "step by step in turn" method which is similar to cases 1, 2. This is possible by using methods of case 1 and that of case 2 alternatively. It will be also possible to treat very complex phenomena efficiently by this method to quantify intercorrelated patterns, corresponding to our purpose to predict events in a valid sense.

I express my gratitude to Professor Z. Suetuna and my colleague Mr. H. Akaike.

## References

[1] C. Hayashi: On the quantification of qualitative data from the mathematico-statistical point of view. Annals of the Institute of Statistical Mathematics, **2**, No. 1 (1950).

[2] C. Hayashi: On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematico-statistical point of view. Annals of the Institute of Statistical Mathematics, **3**, No. 2 (1952).

[3] C. Hayashi: Multidimensional quantification, I. Proceedings of the Japan Academy, **30**, No. 2 (1954).

[4] C. Hayashi: Multidimensional quantification—with applications to the analysis of social phenomena. Annals of the Institute of Statistical Mathematics, **4**, No. 2 (1954) (in press).

[5] C. Hayashi: Measurement and quantification of social attitude (in Jap.). Proceedings of the Institute of Statistical Mathematics, **1**, No. 2 (1954) (in press).

[6] C. Hayashi: Examples of the theory of quantification (in Jap.). Proceedings of the Institute of Statistical Mathematics, **2**, No. 1 (1954) (in press).