Further Corroboration of Equivalence of Two χ^2 Forms*

MO Xiao-Hu[1)]

(Institute of High Energy Physics, CAS, Beijing 100049, China)

Abstract The equivalence of two χ^2 forms for linear function fit is proved. The two forms of χ^2 are applied on a simplified R-value measurement to test the equivalence.

Key words equivalence, χ^2 form, linear function, R-value measurement

1 Introduction

In physics experiments, the covariance matrix is used to construct the χ^2 estimator for correlated data. The matrix is minimized to acquire the best estimates for measured parameters[1]. Usually the experimental data are affected by overall systematic errors, such as the error of luminosity or efficiency in scan experiment. Without loosing generality, let n_i be the number of selected events for certain final state at the i-th energy point, and ε the corresponding efficiency for all measured points[2)], the number of measured events is calculated as

$$y_i = \frac{n_i}{\varepsilon} \ .$$

Since all y_i's contain the same ε, they are correlated and ε here could be treated as a normalization factor. Under such circumstances, the $n \times n$ covariance matrix V for n measurements are to be constructed as follows: the diagonal elements are given by

$$v_{ii} = \sigma_i^2 + y_i^2 \sigma_{\mathrm{f}}^2 \ ,$$

where σ_i is the statistic uncertainty of n_i, σ_{f} is the relative uncertainty of ε and the quantity $y_i\sigma_{\mathrm{f}}$ is the normalization uncertainty due to the factor ε for the variable y_i. The correlation between data points i and j contributes to the off-diagonal matrix element v_{ij}, which is the product of two normalization uncertainties, i.e.

$$v_{ij} = y_i y_j \sigma_{\mathrm{f}}^2 \ .$$

The convariance matrix is expressed explicitly as[3)]

$$V = \begin{pmatrix} \sigma_1^2 + y_1^2\sigma_{\mathrm{f}}^2 & y_1 y_2 \sigma_{\mathrm{f}}^2 & \cdots & y_1 y_n \sigma_{\mathrm{f}}^2 \\ y_2 y_1 \sigma_{\mathrm{f}}^2 & \sigma_2^2 + y_2^2\sigma_{\mathrm{f}}^2 & \cdots & y_2 y_n \sigma_{\mathrm{f}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_n y_1 \sigma_{\mathrm{f}}^2 & y_n y_2 \sigma_{\mathrm{f}}^2 & \cdots & \sigma_n^2 + y_n^2\sigma_{\mathrm{f}}^2 \end{pmatrix} . \tag{1}$$

We want to minimize[4)]

$$\chi^2_{\mathrm{M}} = \eta^{\tau} V^{-1} \eta, \tag{2}$$

where

$$\eta = \begin{pmatrix} y_1 - k_1 \\ y_2 - k_2 \\ \vdots \\ y_n - k_n \end{pmatrix}$$

is the vector of the residuals between experimental observations y_i and theoretical expectation k_i.

Apart from the matrix method, another alternative way to handle correlation is the so-called factor method[3, 4]. In this method, a normalization factor

Received 6 June 2005, Revised 28 July 2005

*Supported by National Natural Science Foundation of China (10491303) and 100 Talents Programme of CAS (U-25)

1) E-mail: moxh@mail.ihep.ac.cn

2)This is a good approximation for the efficiency at continuum region, where no resonance exists.

3)The convariance matrix is derived formally by the error propagation forumla with details in Ref. [5].

4)In this paper, chi-square minimization is adopted to obtain the best estimated value. For experimental data minimization, the MINUIT package is used. More on chi-square minimization technique and MINUIT package are found in Ref. [2].

f is introduced and to be fitted as a free parameter to take the correlation into account

$$\chi^2_{\rm f} = \sum_i \frac{(fy_i - k_i)^2}{\sigma_i^2} + \frac{(f-1)^2}{\sigma_{\rm f}^2} \ . \tag{3}$$

where σ_i and $\sigma_{\rm f}$ have the same meaning as in the matrix method clarified before.

Actually, we can comprehend Eq. (3) from an experimental point of view. We treat the efficiency ε also as a measured variable and introduce a fitting parameter ε_0 as its expected value. Then we construct a chisquare form for $n+1$ uncorrelated measurements as follows:

$$\chi^2 = \sum_i \frac{(y_i' - k_i)^2}{\sigma_i^2} + \frac{(\varepsilon - \varepsilon_0)^2}{(\delta\varepsilon)^2} \ , \tag{4}$$

where $\delta\varepsilon$ is the uncertainty on efficiency and $y_i' \equiv n_i/\varepsilon_0$. The latter can be rewritten as

$$y_i' = \frac{n_i}{\varepsilon_0} = \frac{n_i}{\varepsilon}\frac{\varepsilon}{\varepsilon_0} = fy_i \ .$$

Here we define $f \equiv \varepsilon/\varepsilon_0$. At the same time, we change the form of the last term in Eq. (4), viz.

$$\frac{(\varepsilon/\varepsilon_0 - \varepsilon_0/\varepsilon_0)^2}{(\delta\varepsilon/\varepsilon_0)^2} = \frac{(f-1)^2}{(\delta\varepsilon/\varepsilon_0)^2} \ .$$

Exprimentally, ε_0 could be replaced by the measurable quantity ε, so $\delta\varepsilon/\varepsilon_0 \simeq \delta\varepsilon/\varepsilon = \sigma_{\rm f}$. Thus we immediately recover Eq. (3).

Now there are two χ^2 forms, as expressed in Eqs. (2) and (3) respectively, which treat the correlated data. As a matter of fact, the equivalence between these two χ^2 forms was first discussed by D'Agostini in Ref. [5]. At the same time, the author pointed out the biasness of these two estimators. Nevertheless, D'Agostini's study was restricted to the two-variate and constant fitting case. It is natural to raise the following questions:

1. Is the equivalence of the two χ^2 forms a general conclusion?

2. Does the biasness of two χ^2 forms always exist?

As to the first question, their equivalence has been proved in the case of multi-variate and constant fitting[6, 7]. Recently, efforts have been made to extend the proof to the case of non-constant fitting. In this paper, first, the case of linear dependence is considered, and the equivalence of the two χ^2 forms is proved in the vigor of mathematics. So far as the second question is concerned, the following study indicates that for linear function fit, the biasness still exists. Fortunately, by virtue of factor method, the biasness due to fit can be corrected through the fitted factor.

Besides the experimental requirement, from statistic point of view, the analytic formulas are desirable for it could provide qualitative understanding towards more complicated problems and bring new clues for advanced study. In the following sections, we first derive the minimization starting from the two χ^2 forms, and demonstrate that they lead to exactly the same expression and so are equivalent. Then the two χ^2 forms are applied on some simplified experiments to test their equivalence quantitatively.

2 Proof of equivalence of two χ^2 forms

We will study the two χ^2 forms in Eqs. (2) and (3), where k_i depends on i linearly, i.e.

$$k_i = \alpha \cdot i + \beta \ .$$

However in actual experiments, the physical attribute is often indicated by a physical variable instead of the sequence of the experiments. For the i-th experiment at energy $E_{\rm cm} = x_i$, we usually use x_i instead of i to denote such experiment. So we write the linear form of k_i as

$$k_i = \alpha \cdot x_i + \beta \ . \tag{5}$$

With the above expression, Eq. (2) reads explicitly

$$\chi^2_{\rm M} = \sum_{i=1}^{n}\sum_{j=1}^{n} [y_i - (\alpha x_i + \beta)] \cdot \lambda_{ij} \cdot [y_j - (\alpha x_j + \beta)] \ . \tag{6}$$

Here, the symbol $(V^{-1})_{ij}$ has been changed into λ_{ij} for the convenience of notation. For the factor method, chi-square has the relatively simpler form:

$$\chi^2_{\rm f} = \sum_{i=1}^{n} \frac{[fy_i - (\alpha x_i + \beta)]^2}{\sigma_i^2} + \frac{(f-1)^2}{\sigma_{\rm f}^2} \ . \tag{7}$$

In the above equations, subscripts M and f indicate the matrix and the factor methods respectively.

2.1 Expectation and variance from covariance matrix method

For simplicity, we introduce the matrix notation

$$T=\frac{1}{2}\begin{pmatrix}\dfrac{\partial^2\chi^2_{\rm M}}{\partial\alpha\partial\alpha} & \dfrac{\partial^2\chi^2_{\rm M}}{\partial\alpha\partial\beta}\\ \dfrac{\partial^2\chi^2_{\rm M}}{\partial\beta\partial\alpha} & \dfrac{\partial^2\chi^2_{\rm M}}{\partial\beta\partial\beta}\end{pmatrix}\equiv\begin{pmatrix}T_{xx} & T_x\\ T_x & T_0\end{pmatrix} .$$

With the minimization condition

$$\begin{cases}\dfrac{\partial\chi^2_{\rm M}}{\partial\alpha}=0,\\ \dfrac{\partial\chi^2_{\rm M}}{\partial\beta}=0,\end{cases}$$

we have

$$T\begin{pmatrix}\hat\alpha\\ \hat\beta\end{pmatrix}=\begin{pmatrix}T_{xy}\\ T_y\end{pmatrix} . \tag{8}$$

In the above equations, we have defined

$$T_{xx}=\sum_{ij}x_i\lambda_{ij}x_j,\quad T_x=\sum_{ij}x_i\lambda_{ij}=\sum_{ij}\lambda_{ij}x_j,$$
$$T_{xy}=\sum_{ij}x_i\lambda_{ij}y_j,\ T_y=\sum_{ij}\lambda_{ij}y_j,\ \text{ and }\ T_0=\sum_{ij}\lambda_{ij}. \tag{9}$$

Solving Eq. (8), we obtain

$$\begin{pmatrix}\hat\alpha\\ \hat\beta\end{pmatrix}=T^{-1}\begin{pmatrix}T_{xy}\\ T_y\end{pmatrix}=\frac{1}{D_{\rm T}}\cdot\begin{pmatrix}T_0T_{xy}-T_xT_y\\ T_{xx}T_y-T_xT_{xy}\end{pmatrix} ,$$

and covariance of α and β

$$\begin{pmatrix}\sigma^2_{\hat\alpha}\\ \sigma^2_{\hat\beta}\end{pmatrix}=\frac{1}{D_{\rm T}}\cdot\begin{pmatrix}T_0\\ T_{xx}\end{pmatrix} ,$$

with

$$D_{\rm T}\equiv|T|=T_{xx}T_0-T_xT_x . \tag{10}$$

With the variables defined in Eqs. (9) and (10), as well as by formulas (A2) and (A3) in appendix, we yield the following results

$$\begin{cases}\hat\alpha=\displaystyle\sum_{ij}\frac{x_iy_i-x_iy_j}{\sigma_i^2\sigma_j^2}\Big/(S\cdot D_{\rm T}) ,\\ \hat\beta=\displaystyle\sum_{ij}\frac{x_i^2y_j-x_ix_jy_j}{\sigma_i^2\sigma_j^2}\Big/(S\cdot D_{\rm T}) ;\end{cases} \tag{11}$$

and

$$\begin{cases}\sigma^2_{\hat\alpha}=\left(\displaystyle\sum_i\frac{1}{\sigma_i^2}+\sigma_{\rm f}^2\cdot\sum_{ij}\frac{y_i^2-y_iy_j}{\sigma_i^2\sigma_j^2}\right)\Big/(S\cdot D_{\rm T}) ,\\ \sigma^2_{\hat\beta}=\left(\displaystyle\sum_i\frac{x_i^2}{\sigma_i^2}+\sigma_{\rm f}^2\cdot\sum_{ij}\frac{x_i^2y_j^2-x_iy_ix_jy_j}{\sigma_i^2\sigma_j^2}\right)\Big/(S\cdot D_{\rm T}) ,\end{cases} \tag{12}$$

where

$$D_{\rm T}=\frac{1}{S}\cdot\left[\sum_{ij}\frac{x_i^2-x_ix_j}{\sigma_i^2\sigma_j^2}+\sigma_{\rm f}^2\cdot\sum_{ijk}\frac{(x_i^2-x_ix_j)y_k^2-x_i^2y_jy_k-x_iy_ix_jy_j+2x_iy_ix_jy_k}{\sigma_i^2\sigma_j^2\sigma_k^2}\right]. \tag{13}$$

2.2 Expectation and variance from factor method

For simplicity, we introduce the matrix notation

$$\varLambda=\frac{1}{2}\begin{pmatrix}\dfrac{\partial^2\chi^2_{\rm f}}{\partial\alpha\partial\alpha} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial\alpha\partial\beta} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial\alpha\partial f}\\ \dfrac{\partial^2\chi^2_{\rm f}}{\partial\beta\partial\alpha} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial\beta\partial\beta} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial\beta\partial f}\\ \dfrac{\partial^2\chi^2_{\rm f}}{\partial f\partial\alpha} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial f\partial\beta} & \dfrac{\partial^2\chi^2_{\rm f}}{\partial f\partial f}\end{pmatrix}\equiv\begin{pmatrix}\mathscr{A} & \mathscr{D} & -\mathscr{E}\\ \mathscr{D} & \mathscr{B} & -\mathscr{F}\\ -\mathscr{E} & -\mathscr{F} & \mathscr{C}\end{pmatrix}.$$

According to the minimization condition

$$\begin{cases}\dfrac{\partial\chi^2_{\rm f}}{\partial\alpha}=0,\\ \dfrac{\partial\chi^2_{\rm f}}{\partial\beta}=0,\\ \dfrac{\partial\chi^2_{\rm f}}{\partial f}=0,\end{cases}$$

we have

$$\varLambda\begin{pmatrix}\alpha\\ \beta\\ f\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \varDelta\end{pmatrix} . \tag{14}$$

In the above equations, we used the definitions

$$\mathscr{A}=\sum_i\frac{x_i^2}{\sigma_i^2},\ \mathscr{B}=\sum_i\frac{1}{\sigma_i^2},\ \ \mathscr{C}=\frac{1}{\sigma_{\rm f}^2}+\sum_i\frac{y_i^2}{\sigma_i^2} ,$$
$$\mathscr{D}=\sum_i\frac{x_i}{\sigma_i^2},\ \mathscr{E}=\sum_i\frac{x_iy_i}{\sigma_i^2},\ \mathscr{F}=\sum_i\frac{y_i}{\sigma_i^2},\ \text{ and} \tag{15}$$
$$\varDelta=\frac{1}{\sigma_{\rm f}^2} .$$

Solving Eq. (14), and utilizing the definitions in Eq. (15), we obtain

$$\begin{cases}\hat\alpha=\displaystyle\sum_{ij}\frac{x_iy_i-x_iy_j}{\sigma_i^2\sigma_j^2}\Big/(\sigma_{\rm f}^2\cdot D_\Lambda) ,\\ \hat\beta=\displaystyle\sum_{ij}\frac{x_i^2y_j-x_ix_jy_j}{\sigma_i^2\sigma_j^2}\Big/(\sigma_{\rm f}^2\cdot D_\Lambda) ,\\ \hat f=\displaystyle\sum_{ij}\frac{x_i^2-x_ix_j}{\sigma_i^2\sigma_j^2}\Big/(\sigma_{\rm f}^2\cdot D_\Lambda) ;\end{cases} \tag{16}$$

and the corresponding covariance of the above quantities

$$\begin{cases} \sigma_{\hat{\alpha}}^2 = \left(\sum_i \frac{1}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{y_i^2 - y_i y_j}{\sigma_i^2 \sigma_j^2}\right) \Big/ (\sigma_f^2 \cdot D_\Lambda) \ , \\ \sigma_{\hat{\beta}}^2 = \left(\sum_i \frac{x_i^2}{\sigma_i^2} + \sigma_f^2 \cdot \sum_{ij} \frac{x_i^2 y_j^2 - x_i y_i x_j y_j}{\sigma_i^2 \sigma_j^2}\right) \Big/ (\sigma_f^2 \cdot D_\Lambda), \\ \sigma_{\hat{f}}^2 = \left(\sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2}\right) \Big/ D_\Lambda \ ; \end{cases} \tag{17}$$

where

$$D_\Lambda \equiv |\Lambda| = \mathscr{A}\mathscr{B}\mathscr{C} - \mathscr{C}\mathscr{D}^2 - \mathscr{B}\mathscr{E}^2 - \mathscr{A}\mathscr{F}^2 + 2\mathscr{D}\mathscr{E}\mathscr{F} = \frac{1}{\sigma_f^2} \cdot \left[\sum_{ij} \frac{x_i^2 - x_i x_j}{\sigma_i^2 \sigma_j^2} + \sigma_f^2 \cdot \sum_{ijk} \frac{(x_i^2 - x_i x_j) y_k^2 - x_i^2 y_j y_k - x_i y_i x_j y_j + 2 x_i y_i x_j y_k}{\sigma_i^2 \sigma_j^2 \sigma_k^2}\right]. \tag{18}$$

By virtue of Eqs. (13) and (18), $S \cdot D_T = \sigma_f^2 \cdot D_\Lambda$. Then comparing Eqs. (11) and (12) versus Eqs. (16) and (17), the equivalence of the two methods can be seen directly.

2.3 Discussion

If we denote $f_i(\boldsymbol{r})$ as a function at the i-th point of any measurements $\boldsymbol{r} = \boldsymbol{r}(x, y, z, \cdots)$ with uncertainty σ_i, the weighted average of $f_i(\boldsymbol{r})$ is defined as

$$\overline{f}(\boldsymbol{r}) = \sigma_s^2 \cdot \sum_i \frac{f_i(\boldsymbol{r})}{\sigma_i^2} \ ,$$

with

$$\frac{1}{\sigma_s^2} \equiv \sum_{i=1}^n \frac{1}{\sigma_i^2} \quad \text{or} \quad \sigma_s^2 \equiv 1 \Big/ \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right).$$

Using the above definition, we can write the optimized quantities of Eqs. (16) and (17) in other forms. For example, we have

$$\hat{f} = \left(\frac{1}{\sigma_s^2} \cdot \frac{\overline{x^2}}{\sigma_s^2} - \frac{\overline{x}}{\sigma_s^2} \cdot \frac{\overline{x}}{\sigma_s^2}\right) \Big/ (\sigma_f^2 \cdot D_\Lambda) \ , \tag{19}$$

and

$$\sigma_{\hat{f}}^2 = \sigma_f^2 \cdot \left(\frac{1}{\sigma_s^2} \cdot \frac{\overline{x^2}}{\sigma_s^2} - \frac{\overline{x}}{\sigma_s^2} \cdot \frac{\overline{x}}{\sigma_s^2}\right) \Big/ (\sigma_f^2 \cdot D_\Lambda) \ . \tag{20}$$

It is interesting to notice that from Eqs. (19) and (20), we obtain the following relation

$$\sigma_{\hat{f}}^2 = \hat{f} \cdot \sigma_f^2 \ . \tag{21}$$

Furthermore, taking advantage of the forms as displayed in Eqs. (19) and (20), we acquire the relation

$$\hat{f} \cdot \overline{y} = \hat{\beta} + \hat{\alpha} \cdot \overline{x} \ . \tag{22}$$

At the same time, from Eq. (5), we have

$$\hat{k}_x = \hat{\beta} + \hat{\alpha} \cdot x \ . \tag{23}$$

Here subscript denotes the dependence of $\hat{k}$ on x. So combining with Eq. (22), we get

$$\hat{k}_{\overline{x}} = \hat{f} \cdot \overline{y} \ . \tag{24}$$

Comparing with the results in Ref. [7], we find the relations expressed in Eqs. (21) and (24) hold for both cases of constant and linear fitting.

3 Experimental test

For illustrative purpose, in this section we apply the formulas to a simplified R value measurement.

In high energy physics, R is defined as the ratio of the hadron production cross section via single photon annihilation to the lowest order point-like QED $\mu^+\mu^-$ cross section $\sigma_{pt} = 4\pi\alpha^2/3s$. In the naive quark-parton model it is expressed as $R = 3\sum_q Q_q^2$, where Q_q is the electric charge of each quark flavor, and the summation runs over all the produced flavors. Taking the lowest order QCD correction and the electro-weak effect into consideration, R value would be larger than the naive value (10/3), and the correction term is a slowly varying function of the center-of-mass (C.M.) energy in the region without any resonances. Therefore, R is well approximated by a linear function.

In experiments, many factors must be considered in R value measurement1). As a pedagogical example, here we take a comparatively concise R expression[4]

$$R = \frac{(N - N_{bg})}{L\varepsilon(1+\delta) \cdot \sigma_{pt}} \ ,$$

where N is the number of the multi-hadronic events detected, N_{bg} is the estimated number of background events, L is the integrated luminosity, $\varepsilon(1+\delta)$ is the

1) The R value measurement at BESII is published in Refs. [9, 10] where the detailed calculation about experiment R value could be found.

acceptance for the multi-hadronic events with radiative effect included and $(1+\delta)$ is the radiative correction factor. Table 1 lists R values measured at thirty eight energy points[8]. For a study of data taken at different times at the same C.M. energy, the estimated systematic point-to-point errors are given to be $\pm 3\%$. For the R value used here, the systematic uncertainty in the detection efficiency ($\pm 8\%$), the luminosity measurement ($\pm 6\%$), the event selection procedure ($\pm 2\%$), and the background substraction ($\pm 3\%$) yielded a common systematic error of $\pm 10\%$, which should be considered as the normalization error. Now these thirty eight R values will be used to test the foregoing conclusions. For minimization, the MINUIT package from CERN library[2] is utilized.

Table 1. Values for R[8]. The errors quoted are point-to-point systematic errors.

$E_{\rm cm}$/GeV	R	ΔR	$E_{\rm cm}$/GeV	R	ΔR
5.60	4.08	0.32	6.60	4.50	0.17
5.70	4.09	0.16	6.65	4.25	0.16
5.75	4.12	0.20	6.70	4.63	0.15
5.80	4.13	0.16	6.75	4.38	0.15
5.85	4.13	0.19	6.80	4.44	0.16
5.90	4.09	0.14	6.85	4.50	0.13
5.95	4.17	0.16	6.90	4.41	0.15
6.00	4.17	0.09	6.95	4.23	0.17
6.05	4.16	0.18	7.00	4.10	0.12
6.10	4.04	0.15	7.05	4.31	0.09
6.15	4.34	0.16	7.10	4.32	0.14
6.20	4.05	0.08	7.15	4.29	0.11
6.25	3.96	0.14	7.20	4.27	0.11
6.30	4.27	0.14	7.25	4.39	0.11
6.35	4.47	0.17	7.30	4.29	0.11
6.40	4.31	0.13	7.35	4.33	0.09
6.45	4.23	0.14	7.40	4.46	0.08
6.50	4.40	0.15	7.45	4.51	0.14
6.55	4.66	0.16	7.50	4.18	0.59

In the χ^2 construction, the following substitutions are made:

$$x_i \to E^i_{\rm cm}\ , \quad y_i \to R^i_{\rm exp.}\ , \quad \sigma_i \to \Delta R^i_{\rm exp.}\ ,$$
$$\beta \to R_0\ , \quad \text{and} \quad \alpha \to \eta\ ,$$

so Eqs. (7) and (6) become

$$\chi^2_{\rm f} = \sum_i \frac{[fR^i_{\rm exp.}-(R_0+\eta E^i_{\rm cm})]^2}{(\Delta R^i_{\rm exp.})^2}+\frac{(f-1)^2}{\sigma^2_{\rm f}}\ ,$$
$$\chi^2_{\rm M} = \sum_{ij}[R^i_{\rm exp.}-(R_0+\eta E^i_{\rm cm})]\times (V^{-1})_{ij}(R^j_{\rm exp.}-(R_0+\eta E^j_{\rm cm})]\ ,$$

where $\sigma_{\rm f}=10\%$ is the overall error of the normalization factor f, and the element (v_{ij}) of matrix V reads

$$v_{ij}=\delta_{ij}\cdot(\Delta R^i_{\rm exp.})^2+\sigma^2_{\rm f}\cdot R^i_{\rm exp.}\cdot R^j_{\rm exp.}\ .$$

The fitting results are summarized in Table 2. At the same time, using Eqs. (16) and (17), we compute the corresponding values theoretically, which are also given in Table 2. We can see that the two different methods lead to the same results up to the significant digits listed in the table. In addition, with these values, we can also test the simple relation given in Eq. (21).

Table 2. Experimental fitted and theoretical calculated values of parameters and relevant information.

parameter	matrix method	factor method	theoretical calculation
R_0	2.2895 ± 0.3772	2.2895 ± 0.3772	2.2895 ± 0.3772
η	0.1241 ± 0.0421	0.1241 ± 0.0421	0.1241 ± 0.0421
f	—	0.7282 ± 0.0853	0.7282 ± 0.0853
χ^2/d.o.f	27.18/35	27.18/35	—

The fitted values of R_0 and η in Table 2 demontrate that the two χ^2 forms give consistent results after minimization. Next we turn to another aspect of the two χ^2 forms, i.e. the fit biasness, which has been noticed in previous papers[5,7,11].

Fig. 1 shows the fitting result, where the solid line represents the best fitted k value. It is obvious that the fitted line is far below all data points. With the matrix χ^2 fit, there is no way to correct this deviation. On the contrary, the factor χ^2 fit provides us with a normalization factor f which just manifests the magnitude of the biasness. In fact, what we want to know is the weighted average of experimental data, i.e. $\overline{y}$. This value should not contain the biasness due to common uncertainty. Eq. (24) gives the relation between $\hat{k}$ and $\overline{y}$, from which we obtain $\overline{y}$ by scaling $\hat{k}$ with the factor f. The dashed line in Fig. 1 denotes the expected $\overline{y}$ which is obtained by rescaling k by the normalization factor f. We can see that the re-scaled line fits the experiment data well. After rescaling, $R_0/f=3.14$ which is consistent with the naive expectation $10/3=3.3$.

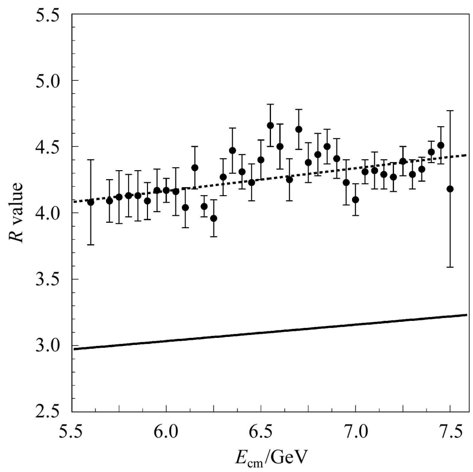

Fig. 1. The R value, error bars indicate point-to-point systematic errors. The data points are taken from Ref. [8]. The solid line represents the best fitted k value, while the dashed line is the expected $\overline{y}$ which is obtained by rescaling k by the normalization factor f.

4 Summary

For linear function fitting, two χ^2 forms have been constructed to handle correlated data. The equivalence of these two forms has been proved in the vigor of mathematics, and tested quantitatively in a simplified experiment which measures R values. However, in light of the comparison of the two χ^2 forms, we notice that the factorized χ^2 form is more transparent. Besides its concise expression and relatively easier minimization, the cause of the biasness is more easily recognized and corrected by the scale factor. By contrast, for matrix χ^2 form, the biasness is uncontrollable.

The importance of this work lies in two aspects. First, the proof of the equivalence of the two χ^2 forms has been extended from constant fitting to linear function fitting; second, the analytical forms of the minimized parameters are given explicitly, which display qualitative features for more complicated fitting.

At last, a remark is in order. Compared with the formulas of the constant fitting, the formulas presented in this paper are more complicated and less intuitive, which implies that further extension of such proof might be more difficult, or an alternative approach should be found.

References

1 Behrend H J et al(CELLO Collab.). Phys. Lett., 1987, **B182**: 400—411

2 CERN Library. MINUIT Reference Manual(Version 92.1). March 1992

3 Brandelik R et al(TASSO Collab.). Phys. Lett., 1982, **B113**: 499—508; Brandelik R et al(TASSO Collab.). Z. Phys., 1980, **C4**: 87—93

4 Bartel W et al(JADE Collab.). Phys. Lett., 1983, **B129**: 145—152

5 D'Agostini G. Nucl. Instrum. Methods, 1994, **A346**: 306—311

6 Takeuchi T. Prog. Thero. Phys., 1996, **123**(Suppl.): 247—264

7 MO Xiao-Hu, ZHU Yong-Sheng. HEP & NP, 2003, **27**: 371—376 (in Chinese)
(莫晓虎, 朱永生. 高能物理与核物理, 2003, **27**: 371—376)

8 Siegrist J L et al. Phys. Rev., 1982, **D26**: 969—990

9 BAI J Z et al. Phys. Rev. Lett., 2000, **84**: 594

10 BAI J Z et al. Phys. Rev. Lett., 2002, **88**: 101802

11 MO Xiao-Hu, ZHU Yong-Sheng. HEP & NP, 2003, **27**: 747—753 (in Chinese)
(莫晓虎, 朱永生. 高能物理与核物理, 2003, **27**: 747—753)

12 Greub Wener. Linear Algebra (Fourth Edition). Springer-Verlag New York Inc., 1981

Appendix A

Matrix formulas

Some formulas on matrix are collected in this appendix, as to the knowledge of matrix, see Ref. [12].

Let's consider a special square matrix A,

$$A=\begin{pmatrix} x_1+\delta & a_1b_2+\delta & \cdots & a_1b_n+\delta \\ a_2b_1+\delta & x_2+\delta & \cdots & a_2b_n+\delta \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1+\delta & a_nb_2+\delta & \cdots & x_n+\delta \end{pmatrix} . \quad \text{(A1)}$$

If its matrix element reads

$$a_{ij}=\delta_{ij}\cdot(x_i-a_ib_i)+a_ib_j+\delta ,$$

then the element of its inverse matrix is expressed as

$$a^{-1}_{ij} = \frac{\delta_{ij}}{x_i-a_ib_i}-\frac{a_ib_j+\delta}{S\cdot(x_i-a_ib_i)\cdot(x_j-a_jb_j)}- \frac{\delta}{S}\cdot\sum_{k=1}^{n}\frac{(a_i-a_k)\cdot(b_j-b_k)}{(x_i-a_ib_i)\cdot(x_j-a_jb_j)\cdot(x_k-a_kb_k)} ,$$

with

$$S=1+\sum_{i=1}^{n}\frac{a_ib_i+\delta}{x_i-a_ib_i}+\delta\cdot\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{a_ib_i-a_ib_j}{(x_i-a_ib_i)\cdot(x_j-a_jb_j)} .$$

In addition, its determinant is

$$\begin{vmatrix} x_1+\delta & a_1b_2+\delta & \cdots & a_1b_n+\delta \\ a_2b_1+\delta & x_2+\delta & \cdots & a_2b_n+\delta \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1+\delta & a_nb_2+\delta & \cdots & x_n+\delta \end{vmatrix} = S\cdot\left[\prod_{i=1}^{n}(x_i-a_ib_i)\right].$$

As a matter of fact, the error matrix given in Eq. (1) is just a special form of the one in Eq. (A1). So with the above formulas, it is easy to acquire the corresponding results of matrix V. If the element of V reads

$$v_{ij}=\delta_{ij}\sigma_i^2+y_iy_j\sigma_{\mathrm{f}}^2 ,$$

then the element of its inverse matrix is

$$\lambda_{ij}=v^{-1}_{ij}=\frac{\delta_{ij}}{\sigma_i^2}-\frac{\sigma_{\mathrm{f}}^2}{S}\cdot\frac{y_iy_j}{\sigma_i^2\sigma_j^2} , \quad \text{(A2)}$$

with

$$S=1+\sigma_{\mathrm{f}}^2\cdot\sum_{i=1}^{n}\frac{y_i^2}{\sigma_i^2} . \quad \text{(A3)}$$

两种χ^2形式等价性证明的推广*

莫晓虎[1)]

(中国科学院高能物理研究所　北京　100049)

摘要　通过直接计算证明了线性函数拟合情况下两种χ^2形式的等价性. 利用简化的R值测量定量地检验了等价性的结论.

关键词　等价性　χ^2形式　线性函数　R值测量

2005 – 06 – 06 收稿, 2005 – 07 – 28 收修改稿

*国家自然科学基金(10491303)和中国科学院百人计划(U-25)资助

1) E-mail: moxh@mail.ihep.ac.cn