Optimal extraction of unknown structure functions from cross sections

A.A. Babich and I.L. Solovtsov

P. Sukhoi GSTU, Gomel

July-August, 2007

Introduction

The cross sections for neutral and charged current deep inelastic scattering is

$$\frac{d^2\sigma}{dx\ dy} \equiv \sigma_{xy} = \sigma_0 \sum_k Y_k(y; Q^2) W_k(x, Q^2),$$

where $Y_k(y; Q^2)$ are known functions in the Bjorken limit they are y-polynomial. $W_k(x, Q^2)$ are hadronic structure functions. For massless leptons the cross sections is parameterized by three unpolarized structure functions

$$W_k^{unpol}(x, Q^2) \Rightarrow F_1(x, Q^2), \quad F_2(x, Q^2), \quad F_3(x, Q^2);$$

and five polarized functions

$$W_k^{pol}(x, Q^2) \Rightarrow g_1(x, Q^2),\ g_2(x, Q^2),\ g_3(x, Q^2),\ g_4(x, Q^2),\ g_5(x, Q^2).$$

To extract structure functions from deep inelastic experimental data one usually uses the some parameterizations

$$f(x, Q^2) = A_0 x^{A_1} (1-x)^{A_2} (1 + A_3 x^{A_4}),$$

and then one fulfil standard χ^2-fitting procedure.

Method

With statistical point of view the task of extraction unknown functions from the differential cross sections may be formulated by following way. Fixing a x-point and Q^2 we get $W_i(x, Q^2) \Rightarrow c_i$ and have for observable $\mathcal{O}(\phi)$ the decomposition form

$$\mathcal{O}(\phi) = \sum_i c_i f_i(\phi),$$

where ϕ denotes variables for kinematic parameters upon which $\mathcal{O}$ depends on and the $f_i(\phi)$ are known functions. The solution is based on usage the functional set $\{w_i(\phi)\}$ which orthogonal to the set $\{f_i(\phi)\}$. Firstly determine the orthogonal conditions in configuration space Φ as

$$\int w_i(\phi)\ f_j(\phi)\ d\phi = \delta_{ij}.$$

Then

$$c_i = \int w_i(\phi)\ \mathcal{O}(\phi)\ d\phi$$

c_i are random values

Method

Number of events registrated in region $d\phi$ of phase space has Poisson distribution.

$$\begin{aligned}
\mathcal{O}(\phi) &= \frac{1}{\mathcal{L}_{eff}} n(\phi)\ d\phi, \\
\mathrm{M}[\mathcal{O}(\phi) d\phi] &= \frac{1}{\mathcal{L}_{eff}} \mathrm{M}[n(\phi) d\phi] = \frac{1}{\mathcal{L}_{eff}} n(\phi)\ d\phi = \mathcal{O}(\phi) d\phi, \\
\mathrm{V}[\mathcal{O}(\phi) d\phi] &= \frac{1}{\mathcal{L}_{eff}^2} \mathrm{V}[n(\phi) d\phi] = \frac{1}{\mathcal{L}_{eff}^2} n(\phi)\ d\phi = \frac{1}{\mathcal{L}_{eff}} \mathcal{O}(\phi) d\phi.
\end{aligned}$$

$\mathcal{L}_{eff}$ is effective luminosity, M[a] is the means of a and V[a] is variance of a.

Covariance matrix for extracting coefficients c_i

$$\mathrm{cov}(c_i c_j) \equiv V_{ij} = \frac{1}{\mathcal{L}_{eff}} \int w_i(\phi) w_j(\phi) \mathcal{O}(\phi) d\phi.$$

Problem of statistical errors optimization formulates as task of minimization of the elements V_{ij} by keeping of orthogonal conditions.

$$\begin{aligned}
\delta V_{ij} = \frac{1}{\mathcal{L}_{eff}} \int \delta(w_i(\phi) w_j(\phi)) \mathcal{O}(\phi) d\phi &= 0, \\
\int \delta w_i(\phi) f_j(\phi) &= 0.
\end{aligned}$$

Method

Consider main steps of solution of the system. Functions $w_i(\phi)$ are taken in the decomposition form:

$$w_i(\phi) = \sum_j Y_{ij}(\phi) f_j(\phi),$$

where $Y_{ij}(\phi)$ are unknown functions. The orthogonality conditions give

$$\sum_j \int \delta Y_{ij}(\phi)\; f_j(\phi)\, f_l(\phi) d\phi = 0.$$

Here $\delta Y_{ij}(\phi)$ are variations of Y_{ij}. Optimization condition gives

$$\sum_{k,l} \left\{ \int \delta Y_{ij} f_k f_l \cdot Y_{jl}(\phi)\mathcal{O}(\phi) d\phi + \int \delta Y_{jl} f_l f_k \cdot Y_{ik}(\phi)\mathcal{O}(\phi) d\phi \right\} = 0.$$

From these equations one concludes that $Y_{jl}(\phi)\mathcal{O}(\phi)$ is independent of ϕ.

$$Y_{jl}(\phi)\mathcal{O}(\phi) = X_{jl} \;\Rightarrow\; Y_{jl}(\phi) = X_{jl} \cdot \frac{1}{\mathcal{O}(\phi)},$$

Method

Solution is

$$w_i(\phi) = \sum_j X_{ij} \cdot \frac{f_j(\phi)}{\mathcal{O}(\phi)}.$$

Elements X_{jl} are determined from orthogonal condition

$$\sum_j X_{ij} \cdot \int \frac{f_i(\phi) f_j(\phi)}{\mathcal{O}(\phi)} d\phi = \delta_{ij}.$$

Introduce matrix

$$M_{ij} = \int \frac{f_i(\phi) f_j(\phi)}{\mathcal{O}(\phi)} d\phi, \qquad X_{ij} = (M^{-1})_{ij}.$$

We find

$$c_i = \sum_j X_{ij} I_j, \quad \text{where} \quad I_j = \int f_j(\phi) d\phi.$$

Method

The optimal covariance matrix $V_{ij}^{\rm opt}$

$$V_{ij}^{\rm opt} = \frac{1}{\mathcal{L}_{\rm eff}}\,(M^{-1})_{ij} = \frac{(M^{-1})_{ij}\;\sigma_{\rm tot}}{N},$$

where $\sigma_{\rm tot}$ are total cross-sections, and N is total number of events which registrated in all kinematical space.

In limit of Gaussian statistics to determine a confidence region of c_i one can use the standard χ^2-criterium

$$\chi^2 = \sum_{ij}(c_i - c_i^{(0)})(c_j - c_j^{(0)})V_{ij}^{-1}$$

$c_i^{(0)}$ are values of c_i which considered as initial parameters.

Binning procedure

On practice the differential cross sections are measured only with help of binning procedure. In results we have the set of cross sections $\{o_i\}$ for the set of bins with errors Δo_i, correspondingly. Introducing the matrix and vectors we obtain

$$\mathcal{O} = \sum c_i F_i,$$

where $\mathcal{O}^{\mathrm{T}} \equiv \underline{o}^{\mathrm{T}} = (o_1, o_2, \ldots, o_k)$ and $F_i^{\mathrm{T}} \equiv \underline{f}_i^{\mathrm{T}} = (f_i^1, f_i^2, \ldots, f_i^k)$. Then the orthogonal conditions may be write as

$$W_i^{\mathrm{T}} F_j = \delta_{ij}, \quad \text{where} \quad W_i^{\mathrm{T}} \equiv \underline{w}_i^{\mathrm{T}} = (w_i^1, w_i^2, \ldots, w_i^k).$$

For coefficients c_i we have

$$c_i = W_i^{\mathrm{T}} \mathcal{O}.$$

Then find covariance of c_i:

$$\begin{aligned}\mathrm{cov}[c_i c_j] \quad &\equiv \quad v_{ij}[c] \equiv < \Delta c_i \Delta c_j >= W_i^{\mathrm{T}} < \Delta\mathcal{O}\Delta\mathcal{O}^{\mathrm{T}} > W_j \Rightarrow \\ &\Rightarrow \quad v_{ij}[c] = W_i^{\mathrm{T}} \; \mathrm{V}[\mathcal{O}] \; W_j.\end{aligned}$$

Binning procedure

The vectors W_i can not be determine unambiguously from the orthogonality condition. Optimization procedure leads to a system

$$\delta v_{ij}[c] = 0,$$
$$F_i^{\mathrm{T}} \delta W_j = 0.$$

By using the symmetry of covariance matrix $\mathsf{V}[\mathcal{O}]$, i.e. $\mathsf{V}^{\mathrm{T}}[\mathcal{O}] = \mathsf{V}[\mathcal{O}]$,

$$\delta v_{ij}[c] = W_j^{\mathrm{T}} \ \mathsf{V}[\mathcal{O}] \ \delta W_i + W_i^{\mathrm{T}} \ \mathsf{V}[\mathcal{O}] \ \delta W_j = 0.$$

These equations are valid if

$$W_i^{\mathrm{T}} \ \mathsf{V}[\mathcal{O}] = \sum_j \lambda_{ij} \ F_j^{\mathrm{T}} \quad \Rightarrow \quad W_i^{\mathrm{T}} = \sum_j \lambda_{ij} F_j^{\mathrm{T}} \ \mathsf{V}^{-1}[\mathcal{O}],$$

where λ_{ij} are unknown coefficients. From the conditions of orthogonality we obtain

$$\sum_j \lambda_{ij} \ F_j^{\mathrm{T}} \mathsf{V}^{-1}[\mathcal{O}] F_l \quad = \quad \sum_j \lambda_{ij} M_{jl} = \delta_{il} \Rightarrow$$
$$\Rightarrow \quad \lambda_{ij} = (M^{-1})_{ij}, \quad \text{where} \quad M_{ij} = F_j^{\mathrm{T}} \ \mathsf{V}^{-1}[\mathcal{O}].$$

Binning procedure

Thus

$$W_i^{\mathrm{T}} = \sum (M^{-1})_{ij} \; \left(F_j^{\mathrm{T}} \mathrm{V}^{-1}[\mathcal{O}] \right)$$

$$c_i \;\; = \;\; \sum_j (M^{-1})_{ij} \; \left(F_j^{\mathrm{T}} \mathrm{V}^{-1}[\mathcal{O}] \; \mathcal{O} \right), \qquad v_{ij}^{\mathrm{opt}}[c] = (M^{-1})_{ij}.$$

For simplicity consider observables $\{o_i\}$ as uncorrelated (on bins) cross sections $\{\sigma_i\}$. Then the covariance matrix of the cross sections will be have diagonal form:

$$V_{ij}[\mathcal{O}] = \delta^2 \sigma_i \; \delta_{ij} = \frac{1}{\mathcal{L}_{\mathrm{eff}}} \; \sigma_i \; \delta_{ij}.$$

The inverse matrix may be written as

$$V_{ij}^{-1}[\mathcal{O}] = \frac{\mathcal{L}_{\mathrm{eff}}}{\sigma_i} \; \delta_{ij}.$$

$$M_{ij} = \mathcal{L}_{\mathrm{eff}} \; \sum_k \frac{f_i^k f_j^k}{\sigma_k} = \mathcal{L}_{\mathrm{eff}} \; m_{ij},$$

where the sum produce on bins.

Binning procedure

$$c_i = \frac{1}{\mathcal{L}_{\text{eff}}} \sum_l m_{il}^{-1} \left(\sum_{k,r} f_l^k \; v_{kr}^{-1}[\mathcal{O}]\sigma_r \right) = \sum_l m_{il}^{-1} \left(\sum_k f_l^k \right)$$

$$v_{ij}^{\text{opt}}[c] = (M^{-1})_{ij} = \frac{1}{\mathcal{L}_{\text{eff}}} m_{ij}^{-1} = \frac{m_{ij}^{-1}\sigma_{\text{tot}}}{N_{\text{tot}}}.$$

In case when values c_i extracted from experimental data the matrix $V_{ij}[\mathcal{O}]$ should be replace on error matrix

$$V_{ij}^{\text{exp}} = (\Delta\sigma_i)^2 \; \delta_{ij}.$$

And we consequently receive

$$m_{ij}^{\text{exp}} = \sum_k \frac{f_i^k f_j^k}{(\Delta\sigma_k)^2},$$

$$c_i = \sum_j (m_{ij}^{\text{exp}})^{-1} \left(\sum_k \frac{f_j^k \sigma_k}{(\Delta\sigma_k)^2} \right), \quad v_{ij}^{\text{opt}}[c] = (m_{ij}^{\text{exp}})^{-1} \cdot \frac{\sigma_{\text{tot}}}{N_{\text{tot}}}.$$

Conclusion

We formulate the method for extracting coefficients c_i from linear decomposition observable $\mathcal{O}$ on set of known functions $\{f_i\}$. The method based on using the orthogonal weight functions $\{w_i\}$. The weight functions $\{w_i\}$ are choosed that to provide minimum for elements of covariance matrix $< \Delta c_i \Delta c_j >$. The method may be applied for extracting hadron structure function from differential cross sections in deep inelastic processes because in Bjorken limit their dependence from kinematical parameters x and y is factorized (see (2)). Should be note for more realistic applying the method should be generalize on case with correlation between bins. Moreover inputs from systematic errors should take to account too.