Notes on Kinematic Fit

Bono, Bono[1]

April 22, 2025

[1] Bono University

Introduction

Assume that we have measured a momentum of two particles, which decay from a mother particle with an exact momentum P_0. In the real world, every measurement inherently carries some errors. Consequently, the measured momentum may not satisfy momentum conservation:

$$P_0 = P_1 + P_2; \quad P_{1,meas} + P_{2,meas} \neq P_0 \tag{1}$$

However, since these measurements did not incorporate our prior knowledge from physics, we can make a more informed estimate of the measured parameters. If the momentum resolution of particles 1 and 2 is well-known, then we can express χ^2 as

$$\chi^2 = \frac{(P_1 - P_{1,meas})^2}{\sigma_1^2} + \frac{(P_2 - P_{2,meas})^2}{\sigma_2^2} \tag{2}$$

Introduction

By incorporating the *Kinematic Constraints*, specifically *momentum conservation*, into our example, we introduce additional terms known as *Lagrange Multiplier* to Equation (3):

$$\chi^2 = \frac{(P_1 - P_{1,meas})^2}{\sigma_1^2} + \frac{(P_2 - P_{2,meas})^2}{\sigma_2^2} + 2\lambda(\mathbf{P_{1,meas}} + \mathbf{P_{2,meas}} - \mathbf{P_0}) \tag{3}$$

We then proceed to evaluate the conditions for local minima, i.e. setting the partial derivatives equal to zero:

$$\frac{1}{2}\frac{\partial \chi 2}{\partial P_1} = \frac{(P_1 - P_{1,meas})}{\sigma_1^2} + \lambda = 0 \tag{4}$$

$$\frac{1}{2}\frac{\partial \chi 2}{\partial P_2} = \frac{(P_2 - P_{2,meas})}{\sigma_2^2} + \lambda = 0 \tag{5}$$

$$\frac{1}{2}\frac{\partial \chi 2}{\partial \lambda} = (P_1 + P_2 - P_0) = 0 \tag{6}$$

Introduction

By solving the equations 4,5,6, we obtain the following expressions:

$$P_1 = \frac{\sigma_2^2 P_{1,meas} - \sigma_1^2 P_{2,meas} + \sigma_1^2 P_0}{\sigma_1^2 + \sigma_2^2} \tag{7}$$

$$P_2 = \frac{\sigma_1^2 P_{2,meas} - \sigma_2^2 P_{1,meas} + \sigma_2^2 P_0}{\sigma_1^2 + \sigma_2^2} \tag{8}$$

$$\lambda = \frac{P_{1,meas} + P_{2,meas} - P_0}{\sigma_1^2 + \sigma_2^2}. \tag{9}$$

Now, we have obtained the 'corrected' measurements with minimized χ^2, which incorporates momentum conservation. Let us delve into the interpretation of these equations.

Introduction

In a straightforward interpretation, λ can be viewed as a kind of 'normalized variance' of the kinematic constraint. It quantifies the error in momentum conservation($P_{1,meas} + P_{2,meas} - P_0$) relative to the overall resolution($\sigma_1^2 + \sigma_2^2$). Equation (4) implies that

$$P_1 = P_{1,meas} - \sigma_1^2 \lambda, \tag{10}$$

suggesting that the corrected momentum(P_1) is essentially the measured momentum($P_{1,meas}$) augmented by a term proportional to the detector resolution and the normalized error of the kinematic constraint. Thus, we can assert that we have applied a statistically fair correction to the momentum, taking into account both the detector resolution and kinematic constraints.

Generalization to Multi-Variables

Assume that you have a set of measurements, $\mathbf{m} = \{m_1, m_2 \dots m_N\}$, and some unmeasured data, $\mathbf{u} = \{u_1, u_2 \dots u_J\}$ to be estimated. Kinematic constraints can be represented by sets of equations $\mathbf{f} = \{f_1(m_1, m_2, \dots m_N, u_1, u_2, \dots u_N), f_2, \dots f_K\}$. We will iteratively solve the problem by guessing the best parameter for each step and checking χ^2. Let $\mathbf{m}^0$ denote our initial measured data, and $\mathbf{m}$ represent the 'guess' of the data in each iterative step.

$$\chi^2(\mathbf{m}) = (\mathbf{m}^0 - \mathbf{m})^\dagger V^{-1}(\mathbf{m}^0 - \mathbf{m}) + 2\lambda^\dagger \mathbf{f}(\mathbf{m},\mathbf{u}). \tag{11}$$

Here, the Lagrange multiplier $\lambda = \{\lambda_1, \lambda_2, \dots \lambda_K\}$ is not just a number but a column vector with k elements, corresponding to each kinematic constraint in $\mathbf{f}$. Our task is to minimize χ^2 to obtain the best guesses in statistically fair method.

χ^2 Minimization

By (partially)differentiating with respect to all variables involved, we obtain the gradients of χ^2. Setting all of them to zero indicates that we have reached a minimum point of χ^2. We have 3 sets of gradient equations:

$$\nabla_{\mathbf{m}} = -2V^{-1}(\mathbf{m}^0)(\mathbf{m}^0 - \mathbf{m}) + 2\mathbf{F}^{\dagger}_{\mathbf{m}}(\mathbf{m}, \mathbf{u})\lambda = 0 \tag{12}$$

$$\nabla_{\mathbf{u}} = 2\mathbf{F}^{\dagger}_{\mathbf{u}}(\mathbf{m}, \mathbf{u})\lambda = 0 \tag{13}$$

$$\nabla_{\lambda} = \mathbf{f}(\mathbf{m}, \mathbf{u}). \tag{14}$$

Here, the subscripts denote partial derivatives. i.e. $((\mathbf{F}_m)_{ki} \equiv \frac{\partial f_k}{\partial m_i})$.

Processing Iterative Steps.

We can express the following equations based on the ones provided above:

$$V^{-1}(\mathbf{m}^0)(\mathbf{m}^{\nu+1} - \mathbf{m}^0) + (\mathbf{F}_m^\dagger)^\nu \lambda^{\nu+1} = 0 \tag{15}$$

$$(\mathbf{F}_u^\dagger)^\nu \lambda^{\nu+1} = 0 \tag{16}$$

$$\mathbf{f}^\nu + \mathbf{F}_m^\nu(\mathbf{m}^{\nu+1} - \mathbf{m}^\nu) + \mathbf{F}_u^\nu(\mathbf{u}^{\nu+1} - \mathbf{u}^\nu) = 0. \tag{17}$$

Equation (17) is not a direct consequence of Equation (13) but rather a linear approximation to proceed with our iteration step. Expanding the ∇_λ term with a Taylor series leads to this equation. Note that, as our parameters $\mathbf{m}$ and $\mathbf{u}$ are updated during the step, our constraint matrix $\mathbf{f}$ should also be updated during the iteration. Here, λ should be indexed as $\nu + 1$ since it is a parameter to be guessed in the next step.

Solving the Equation(1)

Multiplying **V** to Equation (15) leads to:

$$\mathbf{m}^{\nu+1} - \mathbf{m}^0 = -V(\mathbf{m}^0)(\mathbf{F}_m^\dagger)^\nu \lambda^{\nu+1}. \tag{18}$$

Substituting Equation (18) into Equation (17), we get:

$$\begin{aligned} \mathbf{F}_u^\nu(\mathbf{u}^{\nu+1} - \mathbf{u}^\nu) &= -\mathbf{f}^\nu - \mathbf{F}_m^\nu(-V(\mathbf{m}^0)(\mathbf{F}_m^\dagger)^\nu \lambda^{\nu+1} + \mathbf{m}^0 - \mathbf{m}^\nu) \\ &= S\lambda^{\nu+1} - R \end{aligned} \tag{19}$$

where$S \equiv \mathbf{F}_\mathbf{m}^\nu V(\mathbf{m}^0)(\mathbf{F}_\mathbf{m}^\dagger)^\nu$ and $R \equiv \mathbf{f}^\nu + \mathbf{F}_\mathbf{m}^\nu(\mathbf{m}^0 - \mathbf{m}^\nu)$. Multiplying $(\mathbf{F}_u^\dagger)^\nu S^{-1}$ and substituting Equation (16), we get:

$$(\mathbf{F}_u^\dagger)^\nu S^{-1}\mathbf{F}_u^\nu(\mathbf{u}^{\nu+1} - \mathbf{u}^\nu) = \cancelto{0}{(\mathbf{F}_u^\dagger)^\nu \lambda^{\nu+1}} - (\mathbf{F}_u^\dagger)^\nu S^{-1} R. \tag{20}$$

Solving the Equation(2)

Then we naturally obtain:

$$\mathbf{u}^{\nu+1} = \mathbf{u}^{\nu} - ((\mathbf{F}_u^{\dagger})^{\nu} S^{-1} \mathbf{F}_u^{\nu})^{-1} (\mathbf{F}_u^{\dagger})^{\nu} S^{-1} R. \tag{21}$$

and from Equation (19)

$$\lambda^{\nu+1} = S^{-1} (\mathbf{F}_u^{\nu} (\mathbf{u}^{\nu+1} - \mathbf{u}^{\nu}) + R) \tag{22}$$

. For a summary, we have obtained all equations to proceed to the next step. All other matrices in the equation can be calculated from parameters of the current step, and χ^2 can be evaluated from (11) .

$$\begin{cases} \mathbf{u}^{\nu+1} = \mathbf{u}^{\nu} - ((\mathbf{F}_u^{\dagger})^{\nu} S^{-1} \mathbf{F}_u^{\nu})^{-1} (\mathbf{F}_u^{\dagger})^{\nu} S^{-1} R & (21) \\ \lambda^{\nu+1} = S^{-1} (\mathbf{F}_u^{\nu} (\mathbf{u}^{\nu+1} - \mathbf{u}^{\nu}) + R) & (22) \\ \mathbf{m}^{\nu+1} = \mathbf{m}^{0} - V(\mathbf{m}^{0}) (\mathbf{F}_m^{\dagger})^{\nu} \lambda^{\nu+1} & (18) \end{cases}$$

Evolution of the Variance Matrix

Take a look at Eq.(18). We see that $m^{\nu+1}$ is an addition(subtraction) of some parameters to the initially measured data. As we already know the error, i.e. Variance matrix, of initial data, we can estimate how error propagates through the fitting process.

$$V(m) = J_{m,m^0} V(m^0) J^{\dagger}_{m,m^0} \tag{23}$$

We need to calculate the Jacobian,

$$J_{m,m^0(i,j)} = \frac{\partial m_i}{\partial m^0_j} \tag{24}$$

Evolution of the Variance Matrix.

To begin with, let us express Eq (18) in terms of m^0. At the moment we will drop the superscript ν. As $\mathbf{f}(\mathbf{m},\mathbf{u})$is a constant on m^0, $\mathbf{F}_{\mathbf{m}}$ also will be a constant to m^0. Then we only need to consider the derivatives of λ. By substituting (21) ,

$$\lambda = S^{-1}(-\mathbf{F}_u(((\mathbf{F}_u^\dagger)S^{-1}\mathbf{F}_u)^{-1}(\mathbf{F}_u^\dagger)S^{-1}R) + R) \tag{25}$$

and we have

$$R \equiv \mathbf{f} + \mathbf{F}_{\mathbf{m}}(\mathbf{m}^0 - \mathbf{m}) \rightarrow \frac{\partial R}{\partial m^0} = \mathbf{F}_m \tag{26}$$

so that

$$\frac{\partial \lambda}{\partial m^0} = S^{-1}(-\mathbf{F}_u((\mathbf{F}_u^\dagger S^{-1}\mathbf{F}_u)^{-1}\mathbf{F}_u^\dagger S^{-1}\mathbf{F}_m) + \mathbf{F}_m). \tag{27}$$

Evolution of the Variance Matrix

Now define the symmetric matrices $G \equiv \mathbf{F}_{\mathbf{m}}^{\dagger} S^{-1} \mathbf{F}_{\mathbf{m}}$, $U \equiv (\mathbf{F}_u^{\dagger} S^{-1} \mathbf{F}_u)^{-1}$ and $H \equiv \mathbf{F}_m^{\dagger} S^{-1} \mathbf{F}_{\mathbf{u}}$. Then we have expressions for $\frac{\partial \lambda}{\partial m^0}$ hence

$$\begin{aligned} J_{m,m^0} &= I - V(m^0)\mathbf{F}_m^{\dagger} \frac{\partial \lambda}{\partial m^0} = I - V\mathbf{F}_m^{\dagger}(-S^{-1}\mathbf{F}_u U^{-1} H^{\dagger} + S^{-1}\mathbf{F}_m) \\ &= I - V(G - HUH^{\dagger}) \end{aligned} \tag{28}$$

If we let $C = G - HUH^{\dagger}$, we obtain

$$V(m) = J_{m,m^0} V J_{m,m^0}^{\dagger} = V - 2VCV + VCVCV. \tag{29}$$

You might want to neglect higher order term, but please keep 2nd order term at the moment. Some materials like [1] had an error in this part.

Pull distribution

It is better to check the pull distribution to check the quality of a fit. By defining the residual $\epsilon = m - m^0$ and its variance $V(\epsilon)$, pull is defined as:,

$$P(\epsilon) = \epsilon / \sqrt{V(\epsilon)} \tag{30}$$

and

$$V(\epsilon) \equiv V(m) + V(m^0) - 2Cov(m, m^0). \tag{31}$$

We have already calculated J_{m,m^0}. Then we directly get the covariance matrix.

Pull distribution

The covariance can be estimated as:

$$Cov(m, m^0) = J_{m,m^0} V(m) = V - VCV. \tag{32}$$

If we substitute this and Eq.(29) into Eq.(31), we get

$$V(\epsilon) = VCVCV. \tag{33}$$

Note that 2nd order term affects the variance.

Variance of the Unknowns

Just like how we derived Eq.(29) we can estimate the variance matrix of the unknowns.

$$V_U = J_{u,m0} V J_{u,m0}^T \tag{34}$$

$J_{u,m0}$ can be obtained from Eq.(21). Denoting $((\mathbf{F}_u^\dagger)^\nu S^{-1} \mathbf{F}_u^\nu)^{-1} (\mathbf{F}_u^\dagger)^\nu S^{-1}$ as K,

$$J_{u^{\nu+1},m^0} = \frac{\partial u^{\nu+1}}{\partial m^0} = \frac{\partial u^\nu}{\partial m^0} - K \frac{\partial R}{\partial m^0} \simeq -K\mathbf{F}_m. \tag{35}$$

Note that we only have initial "Guess" for the unknowns; In principle, it is not a driven value from measurements. Then, $\frac{\partial u^0}{\partial m_0} = 0$. Also, we approximate that the terms in 2nd or higher iterations are negligible: $J_{u,m^0} \simeq J_{u^0,m^0}$.

Example: Mass-Constraint Fit

Assume a decay of $\Lambda \to p\pi^-$. Momentum of the involved particles are represented in spherical coordinate as:

$$\vec{P}_\Lambda = (P_\Lambda, \theta_\Lambda, \phi_\lambda), \cdots \tag{36}$$

Then we define the energy-momentum constraint equation as:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} -P_\Lambda \sin\theta_\Lambda \cos\phi_\Lambda + P_p \sin\theta_p \cos\phi_p + P_\pi \sin\theta_\pi \cos\phi_\pi \\ -P_\Lambda \sin\theta_\Lambda \sin\phi_\Lambda + P_p \sin\theta_p \sin\phi_p + P_\pi \sin\theta_\pi \sin\phi_\pi \\ -P_\Lambda \cos\theta_\Lambda + P_p \cos\theta_p + p_\pi \cos\theta_\pi \\ -\sqrt{P_\Lambda^2 + m_\Lambda^2} + \sqrt{P_p^2 + m_p^2} + \sqrt{P_\pi^2 + m_\pi^2} \end{pmatrix}. \tag{37}$$

where the mass constraint is naturally implemented in energy term.

Example: Mass-Constraint Fit

We have unmeasured and measured variables as:

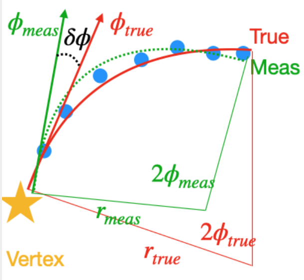
$$\mathbf{u} = \{P_\Lambda, \theta_\Lambda, \phi_\Lambda\}; \quad \mathbf{m} = \{P_p, \theta_p, \phi_p, P_\pi, \theta_\pi, \phi_\pi\}. \tag{38}$$

Since we have 3 unmeasured variable with 4 kinematical constraints, this is a 4-3 = 1-Constrained fit. Let us substitute Eq (37) and (38) into Eq (11)s and its resulting equations. We get $\mathbf{F_u}$ and $\mathbf{F_m}$ as

$$\mathbf{F_u} = \begin{pmatrix} \frac{\partial f_1}{\partial P_\Lambda} & \frac{\partial f_1}{\partial \theta_\Lambda} & \frac{\partial f_1}{\partial \phi_\Lambda} \\ \frac{\partial f_2}{\partial P_\Lambda} & \frac{\partial f_2}{\partial \theta_\Lambda} & \frac{\partial f_2}{\partial \phi_\Lambda} \\ \frac{\partial f_3}{\partial P_\Lambda} & \frac{\partial f_3}{\partial \theta_\Lambda} & \frac{\partial f_3}{\partial \phi_\Lambda} \\ \frac{\partial f_4}{\partial P_\Lambda} & \frac{\partial f_4}{\partial \theta_\Lambda} & \frac{\partial f_4}{\partial \phi_\Lambda} \end{pmatrix}; \quad \mathbf{F_m} = \begin{pmatrix} \frac{\partial f_1}{\partial P_p} & \cdots & \frac{\partial f_1}{\partial \phi_\pi} \\ \frac{\partial f_2}{\partial P_p} & \cdots & \frac{\partial f_2}{\partial \phi_\pi} \\ \frac{\partial f_3}{\partial P_p} & \cdots & \frac{\partial f_3}{\partial \phi_\pi} \\ \frac{\partial f_4}{\partial P_p} & \cdots & \frac{\partial f_4}{\partial \phi_\pi} \end{pmatrix} \tag{39}$$

We have all the matrices to calculate in each step. By applying an appropriate variance matrix and employing χ^2 selection criteria, we can do kinematic fit for the particles.

Covariance Matrix in Helix Track



The cartoon on the left illustrates the correlation between the helical (or circular) track and the azimuth angle at the intersection point (or decay vertex). Since the helical trajectory is defined by measured spatial points, a deviation in radius does not shift the arc itself but moves the center of the circle. It is also worth noting that the arc length, l, remains nearly constant.

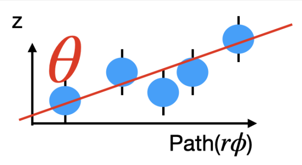
If we let the azimuth angle at the center of the track as ϕ_0, we deduce following relations:

$$\phi = \phi_0 \pm \frac{l}{2r}; \quad \delta\phi = \pm\frac{l}{2r}\frac{\delta r}{r} = \pm\frac{l}{r}\frac{\delta p_T}{p_T} \tag{40}$$

and we naturally obtain:

$$\sigma_\phi^2 = \langle\delta\phi\delta\phi\rangle = \frac{l^2}{r^2 p_T^2}\sigma_{p_T}^2; \quad \mathrm{Cov}(\phi, p_T) = \langle\delta\phi\delta p_T\rangle = \pm\frac{l}{r p_T}\sigma_{p_T}^2 \tag{41}$$

Covariance Matrix in Helix Track



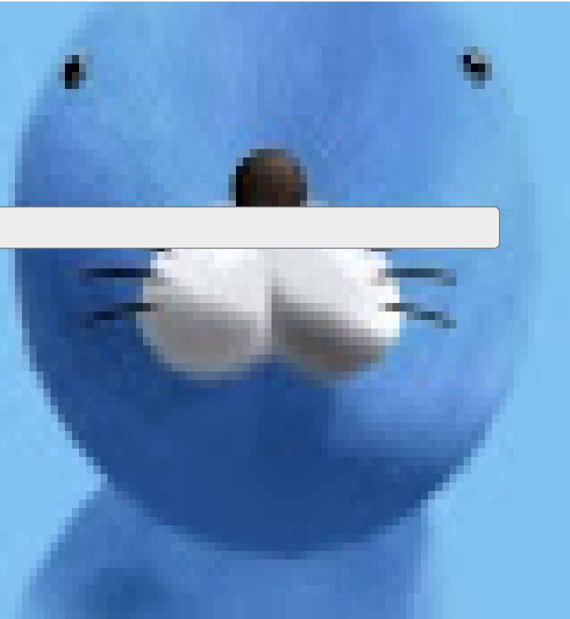
The left figure illustrates the relationship between the pitch(dz) and the helix fit. The pitch is determined by fitting the vertical displacement along transverse path of the helical trajectory. Denoting the polar angle is expressed as $\theta = \frac{\pi}{2} - \arctan(dz)$, we estimate the varaiance of θ based on the fitting error of dz, the slope parameter of linear fit.

$$\sigma_{dz}^2 = \frac{\sum \delta_z^2/(n-2)}{\sum (x-\bar{x})^2} \simeq \frac{n\sigma_z^2/(n-2)}{nL^2/12}; \quad \sigma_\theta = \frac{\partial dz}{\partial \theta}\sigma_{dz} = \frac{1}{1+dz^2}\sigma_{dz}.$$

Note that, the momentum $p_z = p_T dz$ would also have some covariance with θ,

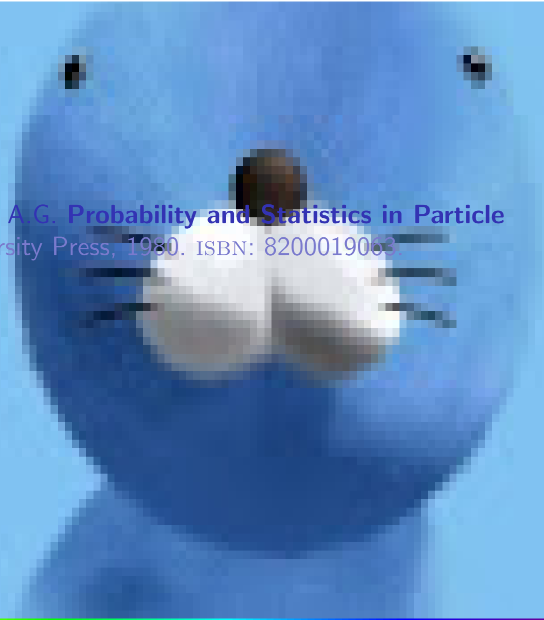
$$\langle \delta p \delta\theta \rangle = dz \cancelto{0}{\langle \delta p_T \delta\theta \rangle} + p_T \langle \delta dz \delta\theta \rangle = \frac{p_T}{1+dz^2}\sigma_{dz}^2.$$

Example: MassVertex-Constraint Fit



To be Updated..

References



O. Skjeggestad Frodesen A.G. **Probability and Statistics in Particle Physics**. Columbia University Press, 1980. ISBN: 8200019063.