# Datenanalyse in der Physik

## Vorlesung 5

## Mehrdimensionale Gauß-Verteilung und Transformation von Wahrscheinlichkeitsdichten

Prof. Dr. J. Mnich

DESY und Universität Hamburg



## Hauptachsentransformation

Erinnerung aus der Linearen Algebra

- Man findet zu jeder symmetrischen $n \times n$ Matrix $A$ eine orthogonale Transformation $U$, d.h $U^{\mathrm{T}} = U^{-1}$, so dass $A' = U\,A\,U^{\mathrm{T}}$ diagonal ist

  $\Longrightarrow$ Diagonalisierung

  Bemerkung: für komplexe Matrizen entspricht das einer unitären Transformation mit $U^\dagger \equiv \left(U^{\mathrm{T}}\right)^* = U^{-1}$

- Das Problem entspricht der Transformation eines Vektors $\vec{y} = U\vec{x}$, so dass die quadratische Form $\vec{y}^{\mathrm{T}}\,A'\,\vec{y}$ nur noch quadratische Terme enthält: $\sum_i A'_{ii}\,y_i^2$

  Hauptachsentransformation (z.B. einer Ellipse in zwei Dimensionen)
  Drehung eines $n$-dimensionenalen Vektors

  Beweis:
  Aus $\vec{y} = U\vec{x}$ folgt $\vec{x} = U^{\mathrm{T}}\vec{y}$ und $\vec{x}^{\mathrm{T}} = \vec{y}^{\mathrm{T}}U$

$$\vec{x}^{\mathrm{T}}\,A\,\vec{x} = \vec{y}^{\mathrm{T}}U\,A\,U^{\mathrm{T}}\vec{y} = \vec{y}^{\mathrm{T}}\,A'\,\vec{y} = \sum_i A'_{ii}y_i^2$$

# Hauptachsentransformation

- Daraus folgt für mehrdimensionale Zufallsvariablen:

  Es ist immer möglich eine orthogonale Transformation $\vec{y} = U\vec{x}$ zu finden, so dass die Zufallsvariablen $y_i$ unkorreliert sind

  d.h. man erhält eine diagonale Kovarianzmatrix und die Korrelationsmatrix wird zur Einheitsmatrix

Beweis:
Wegen $E[\lambda h(x)] = \lambda E[h(x)]$ und $E[h(x) + j(x)] = E[h(x)] + E[j(x)]$ können lineare Transformation auch auf Erwartungswerte angewendet werden

Es gilt für die Kovarianzmatrix der Zufallsvariablen $\vec{x}$

$$V = V(\vec{x}\,) = E[(\vec{x} - \vec{\mu_x})(\vec{x} - \vec{\mu_x})^{\mathrm{T}}]$$

Wir wenden darauf eine orthogonale Transformation $U$ an, so dass $V' = U\,V(\vec{x}\,)\,U^{\mathrm{T}}$ diagonal ist:

$$\begin{aligned} V' &= U\,V(\vec{x}\,)\,U^{\mathrm{T}} = U E[(\vec{x} - \vec{\mu_x})(\vec{x} - \vec{\mu_x})^{\mathrm{T}}]\,U^{\mathrm{T}} \\ &= E[U(\vec{x} - \vec{\mu_x})(\vec{x} - \vec{\mu_x})^{\mathrm{T}}\,U^{\mathrm{T}}] = E[(\vec{y} - \vec{\mu_y})(\vec{y} - \vec{\mu_y})^{\mathrm{T}}] \end{aligned}$$

mit $\vec{y} = U\vec{x}$ und $\vec{\mu_y} = U\vec{\mu_x}$



# Mehrdimensionale Gauß-Verteilung

Die mehrdimensionale Gauß-Verteilung des Zufallsvektors $\vec{x}$ enthält als Parameter den Vektor der Mittelwerte $\vec{\mu_x} = \vec{x}$ und die Kovarianzmatrix $V$

$$\boxed{f(\vec{x}\,) = \frac{1}{(2\,\pi)^{\frac{n}{2}}\,|V|^{\frac{1}{2}}}\,e^{-\frac{1}{2}(\vec{x} - \vec{\mu_x})^{\mathrm{T}}\,V^{-1}\,(\vec{x} - \vec{\mu_x})}}$$

- Darin ist $|V| = \det V$ die Determinante der Kovarianzmatrix
- und $V^{-1}$ die inverse Kovarianzmatrix ($V\,V^{-1} = 1$)
- Die Zahl der Parameter der mehrdimensionalen Gauß-Verteilung ist:
  - $n$ Mittelwerte und
  - $(n^2 + n)/2$ Elemente der symmetrischen Kovarianzmatrix

Das ist das einfachstes und wichtigstes Beispiel einer mehrdimensionalen Wahrscheinlichkeitsdichteverteilung

Wegen der großen Bedeutung wollen wir hier die mehrdimensionale Gauß-Verteilung explizit zeigen

# Beweis der mehrdimensionalen Gauß-Verteilung

Zur Vereinfachung führen wir zunächst folgende Abkürzungen ein

$$f(\vec{x}\,) = N e^{\,g(\vec{x}\,)} \qquad \text{mit} \qquad g(\vec{x}\,) = -\frac{1}{2}\,(\vec{x} - \vec{\mu}_x)^{\mathrm{T}}\, G\,(\vec{x} - \vec{\mu}_x)$$

Wir berechnen den Normierungsfaktor $N$ und zeigen, dass die Matrix $G$ die inverse Kovarianzmatrix ist

Normierung: Es soll gelten

$$1 = \int \ldots \int N e^{\,g(\vec{x}\,)}\,\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Wir führen eine orthogonale Transformation aus $\vec{x} = U\vec{y}$, so dass die Matrix $G' = U^{\mathrm{T}} G\, U$ diagonal wird

$$g(\vec{x}\,) = -\frac{1}{2}\,(\vec{y} - \vec{\mu}_y)^{\mathrm{T}}\, U^{\mathrm{T}} G\, U\,(\vec{y} - \vec{\mu}_y) = -\frac{1}{2}\,(\vec{y} - \vec{\mu}_y)^{\mathrm{T}}\, G'\,(\vec{y} - \vec{\mu}_y) = -\frac{1}{2}\sum_i G'_{ii}\,(y_i - \mu_{y\,i})^2$$

Bei Wechseln der Integrationvaribalen von $\mathrm{d}x_1 \ldots \mathrm{d}x_n$ nach $\mathrm{d}y_1 \ldots \mathrm{d}y_n$ tritt die Funktionaldeterminante auf:

$$\vec{y} = U^{\mathrm{T}}\vec{x} \qquad \Longrightarrow \qquad \frac{\partial y_i}{\partial x_j} = \frac{\partial}{\partial x_j}\sum_k U_{ki} x_k = \sum_k U_{ki}\frac{\partial}{\partial x_j} x_k = U_{ji}$$

Für orthogonale Transformationen gilt aber $\det U = \det U^{\mathrm{T}} = 1$. Damit ist die Normierung

$$1 = \int \ldots \int N e^{\,g(\vec{y}\,)}\,\mathrm{d}y_1 \ldots \mathrm{d}y_n = \int \ldots \int N e^{-\frac{1}{2}\sum_i G'_{ii}\,(y_i - \mu_{y\,i})^2}\,\mathrm{d}y_1 \ldots \mathrm{d}y_n$$



# Beweis der mehrdimensionalen Gauß-Verteilung

Umwandeln der Summe im Exponenten in ein Produkt

$$1 = \int \ldots \int N \prod_i e^{-\frac{1}{2} G'_{ii}\,(y_i - \mu_{y\,i})^2}\,\mathrm{d}y_1 \ldots \mathrm{d}y_n$$

Das Mehrfachintegral zerfällt in ein Produkt von Integralen und mit der Normierung der eindimensionalen Gauß-Verteilung ergibt sich

$$1 = N \prod_i \underbrace{\int e^{-\frac{1}{2} G'_{ii}\,(y_i - \mu_{y\,i})^2}\,\mathrm{d}y_i}_{=\sqrt{\frac{2\,\pi}{G'_{ii}}}} = N \prod_{i=1}^{n} \sqrt{\frac{2\,\pi}{G'_{ii}}} = N \frac{(2\,\pi)^{\frac{n}{2}}}{\sqrt{\prod_{i=1}^{n} G'_{ii}}}$$

Das Produkt $\prod_{i=1}^{n} G'_{ii}$ ist nun aber genau die Determinante der diagonalen Matrix $G'$ und weil eine orthogonale Transformation die Determinante nicht ändert
$\det G' = \det(U^{\mathrm{T}} G\, U) = (\det U^{\mathrm{T}})(\det G)(\det U) = \det G$
folgt

$$N = \frac{\sqrt{\prod_{i=1}^{n} G'_{ii}}}{(2\,\pi)^{\frac{n}{2}}} = \frac{\sqrt{\det G}}{(2\,\pi)^{\frac{n}{2}}} = \frac{1}{(2\,\pi)^{\frac{n}{2}}\,\sqrt{\det(G^{-1})}}$$

# Beweis der mehrdimensionalen Gauß-Verteilung

Wir zeigen nun noch, dass gilt $G = V^{-1}$. Für alle Mittelwerte $\mu_{x\,i}$ soll gelten $E[x_i] = \mu_{x\,i}$ oder

$$0 = E[x_i - \mu_{x\,i}] = \int \ldots \int (x_i - \mu_{x\,i})\, N e^{\,g(\vec{x}\,)}\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Ableiten dieser Gleichung nach dem Parameter $\mu_{x\,l}$ ergibt

$$\begin{aligned} 0 &= \frac{\partial}{\partial \mu_{x\,l}} \int \ldots \int (x_i - \mu_{x\,i})\, N e^{\,g(\vec{x}\,)}\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \\ &= -\delta_{il} \int \ldots \int N e^{\,g(\vec{x}\,)}\, \mathrm{d}x_1 \ldots \mathrm{d}x_n + \int \ldots \int (x_i - \mu_{x\,i})\, N e^{\,g(\vec{x}\,)}\, \frac{\partial g(\vec{x}\,)}{\partial \mu_{x\,l}}\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \end{aligned}$$

worin $\delta_{il} = \begin{cases} 1 & i = l \\ 0 & i \neq l \end{cases}$ und das erste Integral die Normierung ist.

Wir berechnen nun $\frac{\partial g(\vec{x}\,)}{\partial \mu_{x\,l}}$. In Komponenten geschrieben

$$g(\vec{x}\,) = -\frac{1}{2}\,(\vec{x} - \vec{\mu}_x)^{\mathrm{T}}\, G\,(\vec{x} - \vec{\mu}_x) = -\frac{1}{2} \sum_i \sum_j (x_i - \mu_{x\,i})\, G_{ij}\,(x_j - \mu_{x\,j})$$

$$\begin{aligned} \frac{\partial g(\vec{x}\,)}{\partial \mu_{x\,l}} &= -\frac{1}{2} \sum_i \sum_j \frac{\partial}{\partial \mu_{x\,l}} \left[(x_i - \mu_{x\,i})\, G_{ij}\,(x_j - \mu_{x\,j})\right] \\ &= -\frac{1}{2} \sum_i \sum_j \left[\delta_{il}(-1)\, G_{ij}\,(x_j - \mu_{x\,j}) + \delta_{jl}(-1)(x_i - \mu_{x\,i})\, G_{ij}\right] \end{aligned}$$



# Beweis der mehrdimensionalen Gauß-Verteilung

$$\frac{\partial g(\vec{x}\,)}{\partial \mu_{x\,l}} = +\frac{1}{2} \left\{ \sum_j G_{lj}\,(x_j - \mu_{x\,j}) + \sum_i G_{il}\,(x_i - \mu_{x\,i}) \right\}$$

Da $G$ symmetrisch ist, $G_{il} = G_{li}$, sind beide Summen identisch

$$\frac{\partial g(\vec{x}\,)}{\partial \mu_{x\,l}} = \sum_j G_{lj}\,(x_j - \mu_{x\,j})$$

Eingesetzt in $\frac{\partial}{\partial \mu_{x\,l}} E[x_i - \mu_{x\,i}]$ folgt

$$0 = -\delta_{il} + \int \ldots \int (x_i - \mu_{x\,i})\, N e^{\,g(\vec{x}\,)} \sum_j G_{lj}\,(x_j - \mu_{x\,j})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

oder

$$\begin{aligned} \delta_{il} &= \sum_j G_{lj} \underbrace{\int \ldots \int (x_i - \mu_{x\,i})\, f(\vec{x}\,)\,(x_j - \mu_{x\,j})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n}_{= \operatorname{cov}(x_i, x_j) = V_{ij} = V_{ji}} \\ \delta_{il} &= \sum_j G_{lj} V_{ji} \end{aligned}$$

In Matrixschreibweise lautet die letzte Gleichung $1 = G \cdot V$ also $G = V^{-1}$

# Transformation von Wahrscheinlichkeitsdichten

Häufige Problemstellung in der Physik:

Messung einer oder mehrerer Größen mit zugehöriger Unsicherheit (Fehler) aus denen physikalische Parameter, inklusive deren Unsicherheiten, bestimmt werden sollen

$\Longrightarrow$ Schätzung von Parametern

Beispiel: Messung der Periodendauer $T$ und Länge $l$ eines Fadenpendels zur Bestimmung der Erdbeschleunigung $g$

$$\omega = \frac{2\,\pi}{T} = \sqrt{\frac{g}{l}} \qquad \Longrightarrow \qquad g = \left(\frac{2\,\pi}{T}\right)^2 l$$

Mögliche Vorgehensweise: Messe Periodendauer $T$ mehrfach und erhalte pdf (z.B. Gauß-Verteilung), deren Standardabweichung $\sigma$ der Messfehler in $T$ abgeschätzt wird, und schätze Fehler in $l$ ab.

Wie groß ist die Standardabweichung in $g$?

$\Longrightarrow$ Transformation von Wahrscheinlicheitsdichten



# Transformation von Wahrscheinlichkeitsdichten

- Einzelne Zufallsvariable

Gegeben sei die Wahrscheinlichkeitsdichteverteilung $f(x)$ der Zufallsvariablen $x$
Suche die Wahrscheinlichkeitsdichteverteilung $g(y)$ in der Zufallsvariablen $y = y(x)$

Die Wahrscheinlichkeit, die Variable $x$ im Intervall $(x, x + \mathrm{d}x)$ zu finden, soll gleich der Wahrscheinlichkeit sein, die Variable $y$ im Intervall $(y, y + \mathrm{d}y)$ zu finden

$$f(x)\,\mathrm{d}x = g(y)\,\mathrm{d}y \qquad \Longrightarrow \qquad g(y) = f(x(y)) \left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = \frac{f(x(y))}{\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|}$$

Dabei ist $x = x(y)$ die Umkehrfunktion zu $y = y(x)$
Falls die Transformation nicht eineindeutig ist, muß über alle Zweige summiert werden

$$g(y) = \sum_{\text{Zweige}} \frac{f(x(y))}{\left|\frac{\mathrm{d}y}{\mathrm{d}x}\right|}$$

Beispiel: die Transformation $y = x^2$ ist für $(-\infty < x < +\infty)$ nicht eineindeutig

$$g(y) = \frac{1}{2\,|x|}\,(f(x) + f(-x)) = \frac{1}{2\,\sqrt{y}}\,(f(+\sqrt{y}) + f(-\sqrt{y}))$$

# Transformation des Mittelwertes und der Varianz

- Für lineare Transformation $y = ax + b$ gilt exakt:

$$\begin{aligned}\langle y\rangle &= E[y] = E[ax+b] = a\,\langle x\rangle + b\\ \sigma_y^2 &= E[(y-\langle y\rangle)^2] = E[(ax+b-a\langle x\rangle - b)^2] = a^2\,E[(x-\langle x\rangle)^2]\\ &= a^2\,\sigma_x^2\end{aligned}$$

- Näherung für nichtlineare Transformationen:
Man entwickele die Funktion $y = y(x)$ um den Mittelwert $x = \langle x\rangle$

$$\begin{aligned} y(x) &= y(\langle x\rangle) + (x-\langle x\rangle)\,\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=\langle x\rangle} + \frac{1}{2}\,(x-\langle x\rangle)^2\,\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{x=\langle x\rangle} + \ldots\\ \Longrightarrow E[y] &= y(\langle x\rangle) + \underbrace{E[x-\langle x\rangle]}_{=0}\,\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=\langle x\rangle} + \frac{1}{2}\,\underbrace{E[(x-\langle x\rangle)^2]}_{=\sigma_x^2}\,\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{x=\langle x\rangle} + \ldots\\ \Longrightarrow \langle y\rangle &\approx y(\langle x\rangle) + \frac{1}{2}\,\sigma_x^2\,\left.\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\right|_{x=\langle x\rangle}\end{aligned}$$

Oft lässt man auch noch den Term $\propto \sigma_x^2$ weg, da ja bereits $\langle x\rangle$ in einer Einzelmessung eine Unsicherheit von $\sigma_x$ hat



# Transformation des Mittelwertes und der Varianz

- Näherung für die Varianz von $y$:
Entwicklung von $y(x)$ bis zum linearen Glied

$$y(x) \approx y(\langle x\rangle) + (x-\langle x\rangle)\,\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=\langle x\rangle}$$

und für $\langle y\rangle \approx y(\langle x\rangle)$ folgt

$$\begin{aligned}(y-\langle y\rangle)^2 &\approx \left((x-\langle x\rangle)\,\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x=\langle x\rangle}\right)^2\\ \Longrightarrow \sigma_y^2 &= E[(y-\langle y\rangle)^2] \approx \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 E[(x-\langle x\rangle)^2] = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\,\sigma_x^2\end{aligned}$$

Das ist das Gesetz der Fehlerfortpflanzung für eine Zufallsvariable:

$$\boxed{\sigma_y^2 = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\,\sigma_x^2}$$

# Beispiel Fehlerfortpflanzung

Häufig hat man es mit Transformationen der Form $y = a\,x^n$ zu tun

$$\Longrightarrow \frac{\mathrm{d}y}{\mathrm{d}x} = n\,a\,x^{n-1} = n\frac{y}{x}$$

Eingesetzt in das Fehlerfortpflanzungsgesetz folgt

$$\sigma_y^2 = \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 \sigma_x^2 \;=\; \left(n\,\frac{y}{x}\right)^2 \sigma_x^2$$

Wenn man nun den relativen Fehler, d.h. die auf den Mittelwert bezogene Standardabweichung $\frac{\sigma_x}{\langle x\rangle}$ bzw. $\frac{\sigma_y}{\langle y\rangle}$ betrachtet, folgt:

$$\frac{\sigma_y}{\langle y\rangle} = n\,\frac{\sigma_x}{\langle x\rangle}$$

Das bedeutet z.B. wenn man den Radius einer Kugel auf $\frac{\sigma_r}{\langle r\rangle} = 1\%$ kennt, so kennt man ihr Volumen nur auf $\frac{\sigma_V}{\langle V\rangle} = 3\%$



# Lineare mehrdimensionale Transformationen

Wie behandeln jetzt die Transformation von $n$ Zufallsvariablen auf $m$ Zufallsvariablen, d.h. von einem $n$-dim. Vektor $\vec{x}$ auf einen $m$-dim. Vektor $\vec{y}$

Achtung: $n$ und $m$ können i.a. verschieden sein

siehe das Beispiel des Fadenpendels:
$n = 2:\; x_1 = T,\; x_2 = l$ und $m = 1:\; y_1 = g$

- Lineare Transformationen ($\to$ exakte Berechnung möglich)

$$\vec{y} = B\,\vec{x} \qquad B \text{ ist } m \times n \text{ Matrix}$$

- Mittelwerte

$$\vec{\mu}_y = E[\vec{y}] = E[B\,\vec{x}] = B\,E[\vec{x}] = B\,\vec{\mu}_x$$

- Kovarianzmatrix

$$\begin{aligned} V'(\vec{y}\,) &= E[(\vec{y}-\vec{\mu}_y)(\vec{y}-\vec{\mu}_y)^{\mathrm{T}}] \;=\; E[B\,(\vec{x}-\vec{\mu}_x)(\vec{x}-\vec{\mu}_x)^{\mathrm{T}}\,B^{\mathrm{T}}] \\ &= B\,E[(\vec{x}-\vec{\mu}_x)(\vec{x}-\vec{\mu}_x)^{\mathrm{T}}]\,B^{\mathrm{T}} \;=\; B\,V(\vec{x}\,)\,B^{\mathrm{T}} \end{aligned}$$

# Lineare mehrdimensionale Transformationen

- Lineare Transformationen

  Man beachte, dass hier für die Ableitungen gilt:

$$\frac{\partial y_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_k B_{ik}\, x_k = B_{ij}$$

- Für den Spezialfall $m = n$, d.h. $B$ ist quadratisch, kann die transformierte Wahrscheinlichkeitsdichteverteilung berechnet werden:

$$\begin{aligned} \int_{\Delta x} \ldots \int f(\vec{x}\,) \mathrm{d}x_1 \ldots \mathrm{d}x_n &= \int_{\Delta y} \ldots \int g(\vec{y}(\vec{x}\,))\, \mathrm{d}y_1 \ldots \mathrm{d}y_n \\ &= \int_{\Delta x} \ldots \int g(\vec{y}(\vec{x}\,))\, |\det J|\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \end{aligned}$$

  wobei beim Variablenwechsel die Funktional- oder Jacobi-Determinante $\det J$ auftritt: $J_{ij} = \frac{\partial y_i}{\partial x_j}$, die im Falle der linearen Transformation die Transformationsmatrix $B$ ist

$$g(\vec{y}\,) = f(\vec{x}\,) \left| \frac{1}{\det B} \right|$$



# Allgemeine mehrdimensionale Transformationen

- Transformation ist gegeben durch $m$ Gleichungen

$$y_i = y_i(x_1, ..., x_n) = y_i(\vec{x}\,) \qquad i = 1, \ldots, m$$

- Im Sonderfall $n = m$ kann wieder die pdf berechnet werden

$$g(\vec{y}\,) = f(\vec{x}\,) \left| \frac{1}{\det B} \right| \qquad \text{mit} \qquad B = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}$$

- Mittelwerte aus der Entwicklung $\vec{y}(\vec{x})$ um die Mittelwerte $\vec{\mu}_x$ (beliebige $m$, $n$)

$$\begin{aligned} y_i(\vec{x}\,) &= y_i(\vec{\mu}_x) + \sum_k (x_k - \mu_{x\,k}) \left. \frac{\partial y_i}{\partial x_k} \right|_{x_k = \mu_{x\,k}} + \ldots \\ \Longrightarrow \langle y_i(\vec{x}\,) \rangle &\approx y_i(\vec{\mu}_x) \qquad \text{oder} \qquad \vec{\mu}_y = \vec{y}(\vec{\mu}_x) \end{aligned}$$

# Allgemeine mehrdimensionale Transformationen

- Lineare Näherung für die Kovarianzmatrix
  Aus Entwicklung von $y_i(\vec{x})$ folgt

$$y_i(\vec{x}) - \mu_{y\,i} \approx \sum_k (x_k - \mu_{x\,k}) \frac{\partial y_i}{\partial x_k}$$

Das Element $il$ der transformierten Kovarianzmatrix $V'(\vec{y})$ ist dann

$$\begin{aligned} V'_{il} &= E[(y_i - \mu_{y\,i})(y_l - \mu_{y\,l})] \\ &\approx E\left[\left(\sum_k (x_k - \mu_{x\,k}) \frac{\partial y_i}{\partial x_k}\right)\left(\sum_j (x_j - \mu_{x\,j}) \frac{\partial y_l}{\partial x_j}\right)\right] \\ &= E\left[\sum_k \sum_j \underbrace{\frac{\partial y_i}{\partial x_k}}_{=B_{ik}} \underbrace{(x_k - \mu_{x\,k})(x_j - \mu_{x\,j})}_{\mathrm{cov}(x_k, x_j)} \underbrace{\frac{\partial y_l}{\partial x_j}}_{=B_{lj}}\right] \end{aligned}$$

$\Longrightarrow$ Allgemeines Gesetz der Fehlerfortpflanzung

$$\boxed{V'(\vec{y}) = B\, V(\vec{x})\, B^{\mathrm{T}}}$$



# Beispiel Fehlerfortpflanzung

Auf einem Messtisch werden kartesische Koordinaten $x_1, x_2$ mit den Standardabweichungen (Fehlern) $\sigma_{x_1}$, $\sigma_{x_2}$ gemessen. Die beiden Messungen sollen unabhängig sein.

Wie groß ist der Messfehler in Polarkoordinaten $r, \varphi$?

$$\begin{aligned} x_1 &= r\cos\varphi \\ x_2 &= r\sin\varphi \end{aligned} \quad \Longrightarrow \quad \begin{aligned} y_1 &= r = \sqrt{x_1^2 + x_2^2} \\ y_2 &= \varphi = \arctan\frac{x_2}{x_1} \end{aligned}$$

$$B = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ -\frac{x_2}{r^2} & \frac{x_1}{r^2} \end{pmatrix}$$

$$\begin{aligned} V'(r,\varphi) = B\, V(x_1, x_2)\, B^{\mathrm{T}} &= \begin{pmatrix} \frac{x_1}{r} & \frac{x_2}{r} \\ -\frac{x_2}{r^2} & \frac{x_1}{r^2} \end{pmatrix} \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} \frac{x_1}{r} & -\frac{x_2}{r^2} \\ \frac{x_2}{r} & \frac{x_1}{r^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{r^2}(x_1^2\sigma_{x_1}^2 + x_2^2\sigma_{x_2}^2) & \frac{xy}{r^3}(\sigma_{x_2}^2 - \sigma_{x_1}^2) \\ \frac{xy}{r^3}(\sigma_{x_2}^2 - \sigma_{x_1}^2) & \frac{1}{r^4}(x_2^2\sigma_{x_1}^2 + x_1^2\sigma_{x_2}^2) \end{pmatrix} \\ &= \begin{pmatrix} \sigma_r^2 & \mathrm{cov}(r,\varphi) \\ \mathrm{cov}(r,\varphi) & \sigma_\varphi^2 \end{pmatrix} \end{aligned}$$

# Beispiel Fehlerfortpflanzung

Was passiert wenn $m \neq n$ ist?

Wir betrachten unser Beispiel des Fadenpendels mit unkorrelierten Messungen von $T$ und $l$:

$$\begin{aligned} x_1 &= T \\ x_2 &= l \end{aligned} \qquad \Longrightarrow \qquad y_1 = g = \left(\frac{2\,\pi}{T}\right)^2 l$$

Die Matrix der partiellen Ableitungen $B$ ist eine $1 \times 2$ Matrix

$$B = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -2\,\frac{g}{T} & \frac{g}{l} \end{pmatrix}$$

und die neue Kovarianzmatrix ist eine $1 \times 1$ Matrix:

$$V(y_1) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \end{pmatrix} \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix} \begin{pmatrix} \frac{\partial y_1}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} \end{pmatrix} = \left(\frac{\partial y_1}{\partial x_1}\right)^2 \sigma_{x_1}^2 + \left(\frac{\partial y_1}{\partial x_2}\right)^2 \sigma_{x_2}^2$$

Die Beiträge der einzelnen Varianzen addieren sich, gewichtet mit den Quadraten der partiellen Ableitungen

Für unser Fadenpendel findet man also:

$$V(g) = \sigma_g^2 = \left(-2\,\frac{g}{T}\right)^2 \sigma_T^2 + \left(\frac{g}{l}\right)^2 \sigma_l^2$$



# Weiteres Beispiel zur Fehlerfortpflanzung

Die Verallgemeinerung des obigen Beispiels ist das häufige Problem, dass $n$ unabhängige Messungen $x_1, \ldots, x_n$ in die Bestimmung eines Parameter $y$ eingehen $(m = 1)$

Anwendung des allgemeinen Gesetzes führt zu:

$$\begin{aligned} \sigma_y^2 = V(y) &= \begin{pmatrix} \frac{\partial y}{\partial x_1} \cdots \frac{\partial y}{\partial x_n} \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma_n^2 \end{pmatrix} \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix} \\ &= \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i}\right)^2 \sigma_i^2 \end{aligned}$$

Beiträge zum Fehler addieren sich quadratisch, gewichtet mit Ableitungen

Falls gilt $y = c\,x_i^{k_i}$ für alle $i$:

$$\frac{\partial y}{\partial x_i} = c\,k_i\,x_i^{k_i - 1} = k_i\,\frac{y}{x} \qquad \Longrightarrow \qquad \sigma_y^2 = \sum_{i=1}^{n} k_i^2 y^2 \frac{\sigma_i^2}{x_i^2}$$

oder für den relativen Fehler

$$\frac{\sigma_y}{y} = \sqrt{\sum_{i=1}^{n} k_i^2 \frac{\sigma_i^2}{x_i^2}}$$

Relative Fehler addieren sich quadratisch, gewichtet mit Potenzen $k_i$

# Übersicht Transformationsformeln

| Eindimensionale Transformationen $x \to y$ | |
|---|---|
| linear | allgemein |
| $y = a\,x + b$ | $y = y(x)$ |
| $\langle y \rangle = a\,\langle x \rangle + b$ | $\langle y \rangle \approx y(\langle x \rangle)$ |
| $\sigma_y^2 = a^2\,\sigma_x^2$ | $\sigma_y^2 \approx \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 \sigma_x^2$ |
| $g(y) = \frac{1}{a} f(x)$ | $g(y) = f(x) \frac{1}{\left\vert \frac{\mathrm{d}y}{\mathrm{d}x} \right\vert}$ |

| Mehrdimensionale Transformationen $x_1, \ldots, x_n \to y_1, \ldots, y_m$ | |
|---|---|
| linear | allgemein |
| $y_i = \sum_j B_{ij}\,x_j$ | $y_i = y_i(x_1, \ldots, x_n)$ |
| $\langle y_i \rangle = \sum_j B_{ij}\,\langle x_j \rangle$ | $\langle y_i \rangle \approx y_i(\langle x_1 \rangle, \ldots, \langle x_n \rangle)$ |
| $V'(\vec{y}) = B\,V(\vec{x})\,B^{\mathrm{T}}$ | $V'(\vec{y}) \approx B\,V(\vec{x})\,B^{\mathrm{T}} \quad \left(B_{ij} = \frac{\partial y_i}{\partial x_j}\right)$ |
| $g(\vec{y}) = f(\vec{x}) \frac{1}{\vert \det B \vert} \quad (n = m)$ | $g(\vec{y}) = f(\vec{x}) \frac{1}{\vert \det B \vert} \quad (n = m)$ |