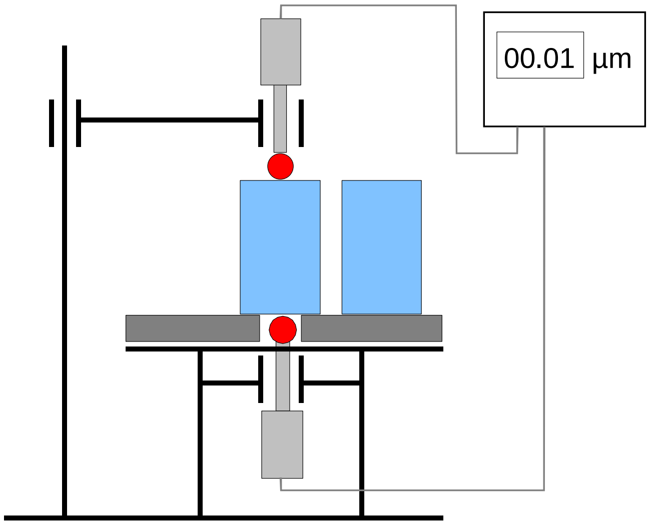
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Institut für Metrologie METAS

Korrelationen

Frédéric Pythoud

Beispiel



- Zwei Endmasse mit Länge (1 et 2) werden durch Vergleich mit derselben Referenz kalibriet:

$$l_1 = l_N + \Delta l_1$$
$$l_2 = l_N + \Delta l_2$$

- Wie geht man mit der Unsicherheit von $l_1 + l_2$ um?

Was ist Korrelation ?

- Gegenteil von „unabhängig“ → abhängig
- Beispiel:
  - Man misst 2 Längen l_1 und l_2 mit dem Referenzendmass
- Problem:
  - Nehmen wir eine Unsicherheit $u_c(l_1)$ für l_1
  - Dann haben wir auch $u_c(l_2)$ für l_1
- Frage
  - Was ist die Unsicherheit auf $l_1 + l_2$ *oder* $l_1 - l_2$

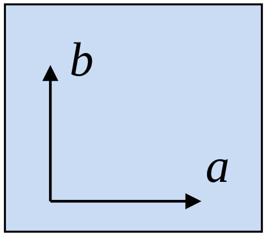
Wie kann man die Korrelation messen ?

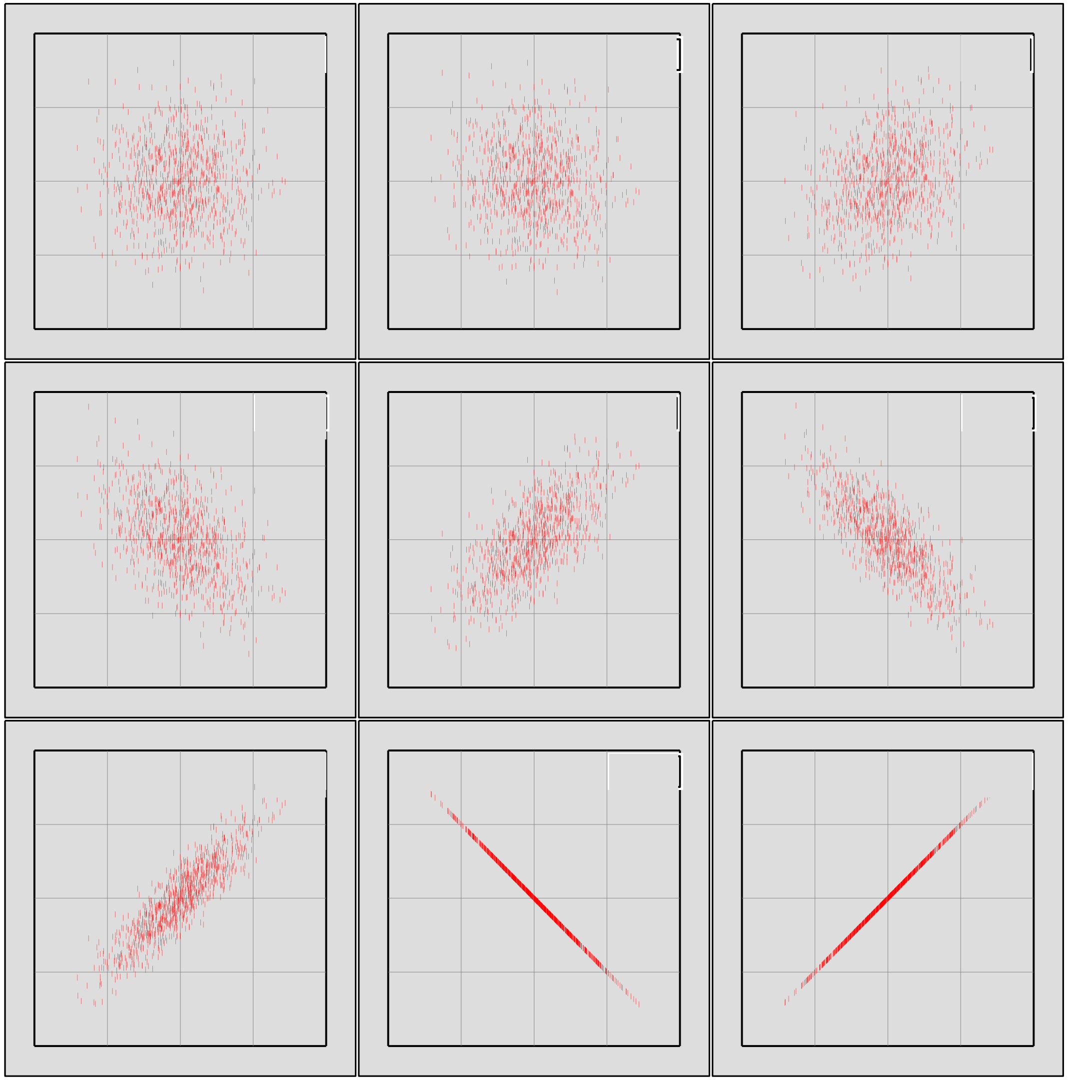
Zwei Messgrössen sind notwendig

- Messgrösse 1: $a_1, a_2, a_3, \dots, a_N$
- Messgrösse 2: $b_1, b_2, b_3, \dots, b_N$

Man kann die Messpunkte $(a_1, b_1), (a_2, b_2)$, in einer Graphik darstellen.

Korrelation graphisch dargestellt





Grafische Repräsentation von zwei gaussverteilten Zufallsgrössen

Jede Grösse hat eine normale Streuung.

Die Korrelation hat erst Bedeutung bei simultaner Betrachtung der Grössen.

Définition der Korrelation

Mittelwerte: $\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$ $\quad \bar{b} = \frac{1}{n}\sum_{i=1}^{n} b_i$

Schätzung der Kovarianz: $s(a,b) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i - \bar{a}) \cdot \left(b_i - \bar{b}\right)$

Varianzen: $s^2(a) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i - \bar{a})^2$ $\quad s^2(b) = \frac{1}{n-1}\sum_{i=1}^{n}(b_i - \bar{b})^2$

Schätzung der Korrelation:
(Definition)

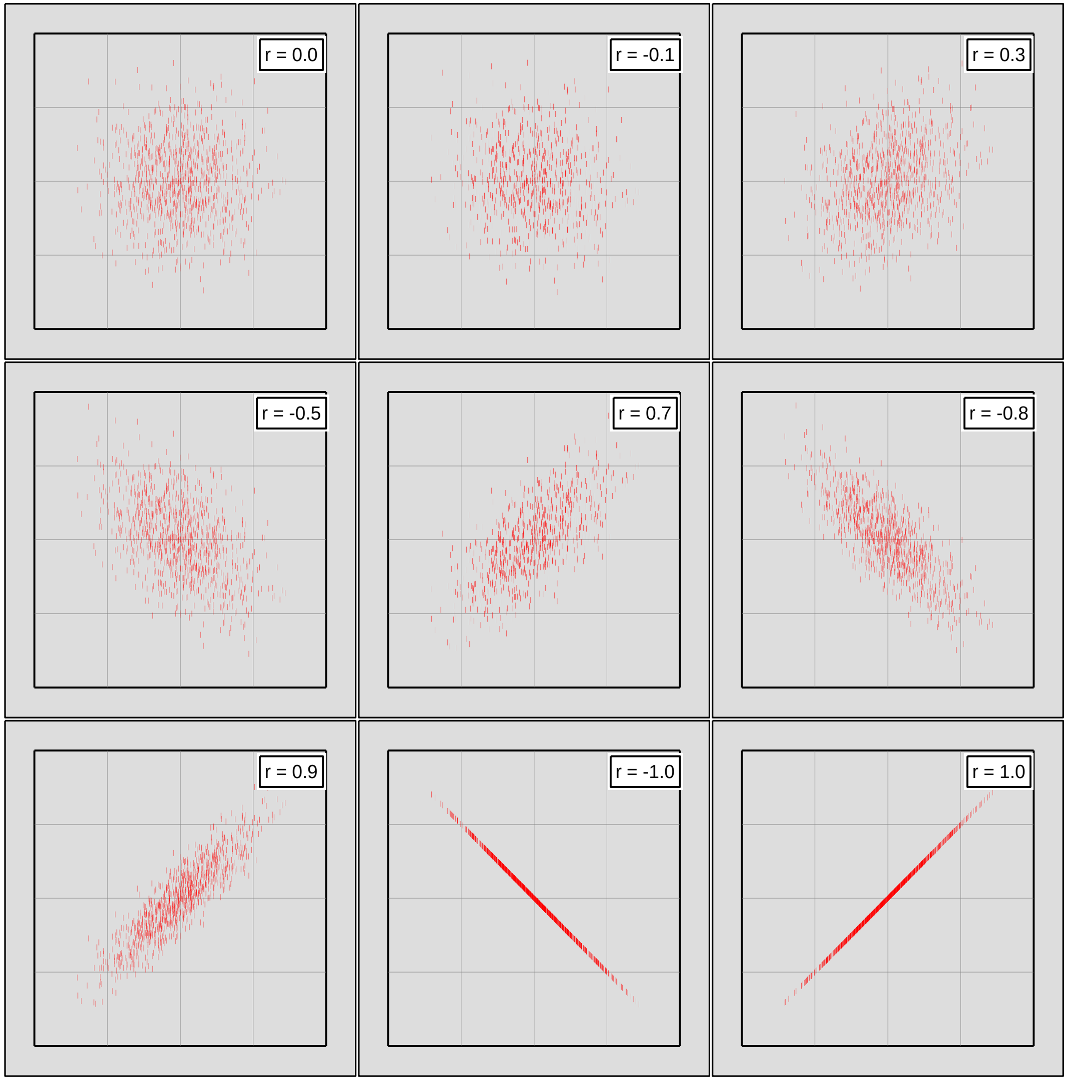
$$r(a,b) := \frac{s(a,b)}{s(a) \cdot s(b)}$$

„Korrelations-
koeffizient“

Eigenschaften der Korrelation

- Das Korrelationskoeffizient hat Werte zwischen -1 und 1.
- Eine Korrelation ungleich null bedeutet Abhängigkeit.
- Korrelation ist intuitiv anschaulicher als Kovarianz.

Korrelation graphisch dargestellt



Unsicherheit des Koeffizients r

Die Schätzung des Korrelationskoeffizienten r ist nicht exakt, sondern mit Unsicherheit behaftet.

Näherung für die erweiterte Unsicherheit ($k = 2$) auf r (abhängig von der Grösse n der Probe und r selbst)

$$r \pm U(r) = \mathrm{arctan}[\left[\frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) \pm \frac{2}{\sqrt{n-3}}\right]$$

- gibt im Allgemeinen ein asymmetrisches Unsicherheitsintervall.
- bei zu kleinem n ist die Berechnung von r sinnlos.

Beispiele

Simultan-Messreihe zweier Messgrössen: n Messpunkte

1) Aus $n = 8$ wird $r = 0.2$ berechnet

Erweitertes Unsicherheitsintervall: $r = 0.2\,^{+0.6}_{-0.8}$

Ohne weiteres mit $r = 0$ verträglich!

2) Aus $n = 8$ wird $r = 0.8$ berechnet

Erweitertes Unsicherheitsintervall: $r = 0.8\,^{+0.16}_{-0.6}$

Korrelation wahrscheinlich, aber Grösse von r immer noch sehr unbestimmt

Formalismus der korrelierten Unsicherheit

Statistik		Physik
Mittelwerte $\bar{a}, \bar{b}$	⇒	Werte der physikalischen Grösse a_0, b_0
Standardabweichung $s(a), s(b)$	⇒	Standardunsicherheit $u(a), u(b)$
Kovarianz $s(a,b)$	⇒	Korrelierte Unsicherheit $u(a,b)$ Bemerkung $: u(a,a) = u^2(a),$ $u(b,b) = u^2(b)$
Korelationskoeffizient $r(a,b) \coloneqq \frac{s(a,b)}{s(a) \cdot s(b)}$	⇒	Korrelationskoeffizient $r(a,b) \coloneqq \frac{u(a,b)}{u(a) \cdot u(b)}$

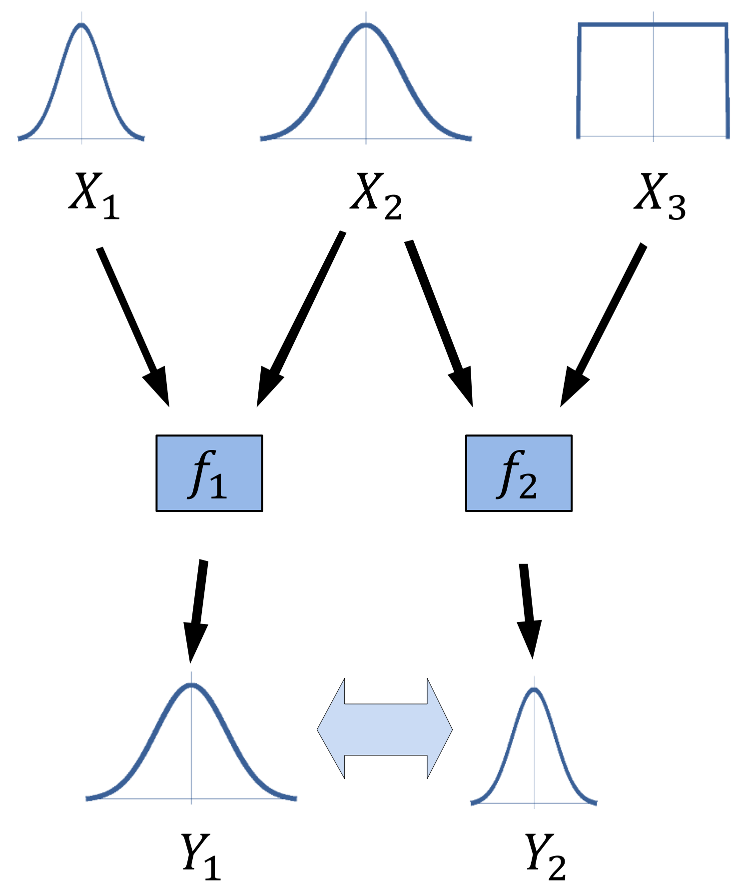
Verwendung der Korrelation

Korrelierte Einflussgrössen

Verallgemeinerung auf das Gesetz der linearen Fortpflanzung von Unsicherheiten

Korrelation zwischen Ausgangsgrössen

- Gemeinsamer Einfluss X_2 führt zu Korrelation zwischen zwei Ausgangsgrössen Y_1 und Y_2
- Frage: Wie gross ist Korrelation?
- **Beispiel:** Messen von 2 Längen mit dem gleichen Normal.



Verallgemeinertes lineares Gesetz zur Fortpflanzung von Unsicherheiten

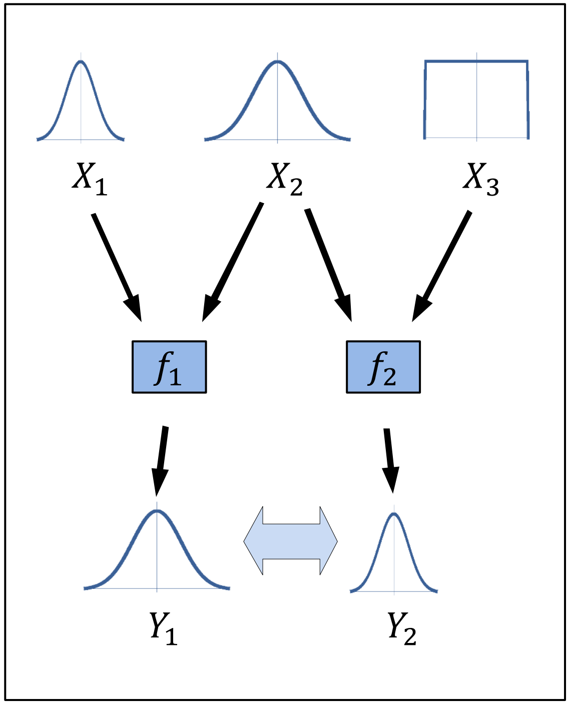
Lineare Fortpflanzung der Unsicherheit

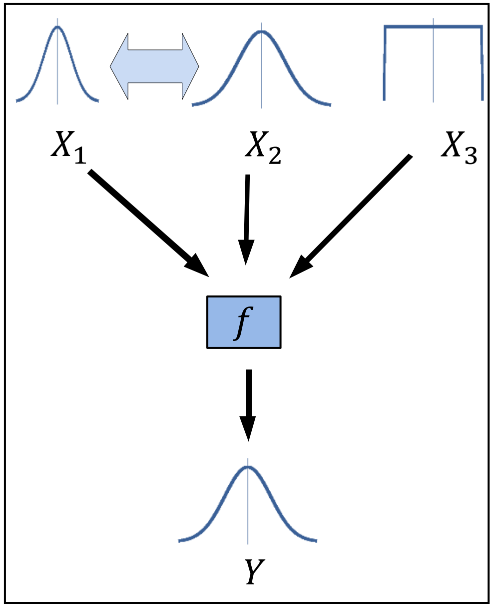
$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i)$$

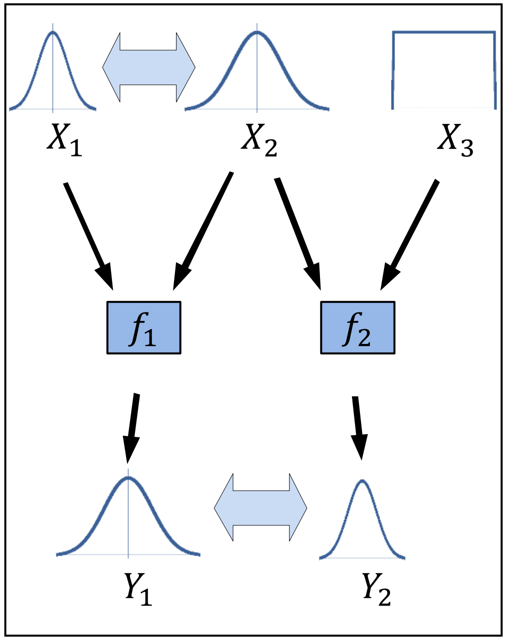
Verallgemeinertes lineares Gesetz zur Fortpflanzung von Unsicherheiten

$$u(y_k, y_l) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial f_k}{\partial x_i} \frac{\partial f_l}{\partial x_j} u(x_i, x_j)$$

Anwendungsbereich







Beispiel Längenmessung

Relativ häufig führen gemeinsam genutzte Standards zu Korrelationen im Messprozess

Mittels eines Längenstandards l_N

- wird Endmass 1 kalibriert. Endmasslänge: $l_1 = l_N + \Delta l_1$
- wird Endmass 2 kalibriert. Endmasslänge: $l_2 = l_N + \Delta l_2$

Beide Endmasse 1 und 2 werden hintereinander gelegt, damit eine totale Länge l bestimmt wird. Die Modellgleichungen sind

$$L_1 = L_N + \Delta L_1$$
$$L_2 = L_N + \Delta L_2$$
$$L = L_1 + L_2$$

Fragen:

- Was ist die Korrelation zwischen l_1 und l_2 ?
- Was ist die Standardunsicherheit von l ?

Korrelation von l_1 und l_2

Es gibt drei Eingangsgrössen: $L_{N,}$ ΔL_1, et ΔL_2

$u(L_1, L_2)$ → Verallgemeinertes lineares Gesetz zur Fortpflanzung von Unsicherheiten

$$= \frac{\partial L_1}{\partial l_N} \cdot \frac{\partial L_2}{\partial l_N} u(l_N, l_N) + \frac{\partial L_1}{\partial l_N} \cdot \frac{\partial L_2}{\partial \Delta l_1} u(l_N, \Delta l_1) + \frac{\partial L_1}{\partial l_N} \cdot \frac{\partial L_2}{\partial \Delta l_2} u(l_N, \Delta l_2)$$
$$+ \frac{\partial L_1}{\partial \Delta l_1} \cdot \frac{\partial L_2}{\partial l_N} u(\Delta l_1, l_N) + \frac{\partial L_1}{\partial \Delta l_1} \cdot \frac{\partial L_2}{\partial \Delta l_1} u(\Delta l_1, \Delta l_1) + \frac{\partial L_1}{\partial \Delta l_1} \cdot \frac{\partial L_2}{\partial \Delta l_2} u(\Delta l_1, \Delta l_2)$$
$$+ \frac{\partial L_1}{\partial \Delta l_2} \cdot \frac{\partial L_2}{\partial l_N} u(\Delta l_2, l_N) + \frac{\partial L_1}{\partial \Delta l_2} \cdot \frac{\partial L_2}{\partial \Delta l_1} u(\Delta l_2, \Delta l_1) + \frac{\partial L_1}{\partial \Delta l_2} \cdot \frac{\partial L_2}{\partial \Delta l_2} u(\Delta l_2, \Delta l_2)$$

$$= 1 \cdot 1 \cdot u^2(l_N) + 1 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 0$$
$$+1 \cdot 1 \cdot 0 + 1 \cdot 0 \cdot 1 + 1 \cdot 1 \cdot 0$$
$$+0 \cdot 1 \cdot 0 + 0 \cdot 0 \cdot 0 + 0 \cdot 1 \cdot 1$$

$$= u^2(l_N)$$

Unsicherheit über die Gesamtlänge

Modell der Kalibrierung:

$$L_1 = L_N + \Delta L_1$$
$$L_2 = L_N + \Delta L_2$$

Standardunsicherheiten:

$$u^2(l_1) = u^2(l_N) + u^2(\Delta l_1)$$
$$u^2(l_2) = u^2(l_N) + u^2(\Delta l_2)$$

„Korrelierte Unsicherheit“:

$$u(l_1, l_2)=u^2(l_N)$$

Die kombinierte Unsicherheit wird berechnet:

$$u^2(l_1 + l_2) = u^2(l_1) + u^2(l_2) + 2\,u(l_1, l_2)$$
$$= u^2(l_N) + \mathrm{u}^2(\Delta l_1) + u^2(l_N) + \mathrm{u}^2(\Delta l_2) + 2\,u^2(l_N)$$
$$=4\,u^2(l_N) + u^2(\Delta l_1) + u^2(\Delta l_2)$$

Ähnliches Beispiel – Andere Vorgehensweise

Durch ein Labor durchgeführte Kalibrierungen von L_1 et L_2 und Messung von L ergibt:

Modell für die Kalibrierung:

$$\boxed{\begin{aligned} L_1 &= L_N + \Delta L_1 \\ L_2 &= L_N + \Delta L_2 \end{aligned}}$$

Modell für die Messung:

$$\boxed{L = L_1 + L_2}$$

Erweitertes Modell:

$$\boxed{\begin{aligned} L &= L_1 + L_2 \\ &= L_N + \Delta L_1 + L_N + \Delta L_1 \\ &= 2L_N + \Delta L_2 + \Delta L_1 \end{aligned}}$$

Lineare Fortpflanzung der Unsicherheiten: $u^2(L) = 4\,u^2(L_N) + u^2(\Delta l_1) + u^2(\Delta l_2)$
(unabhängige Einflussgrössen)

Auswirkung auf Grundoperationen

Addition: $Y = f(X_1, X_2) = X_1 + X_2$

Kombinierte Unsicherheit: $u^2(y) = u^2(x_1) + u^2(x_2) + \underbrace{2\, r(x_1, x_2) u(x_1) u(x_2)}_{2u(x_1, x_2)}$

Ist die Korrelation positiv ($r(x_1, x_2) > 0$), erhöht sich die Unsicherheit.

Subtraktion: $Y = f(X_1, X_2) = X_1 - X_2$

Kombinierte Unsicherheit: $u^2(y) = u^2(x_1) + u^2(x_2) - \underbrace{2\, r(x_1, x_2) u(x_1) u(x_2)}_{2u(x_1, x_2)}$

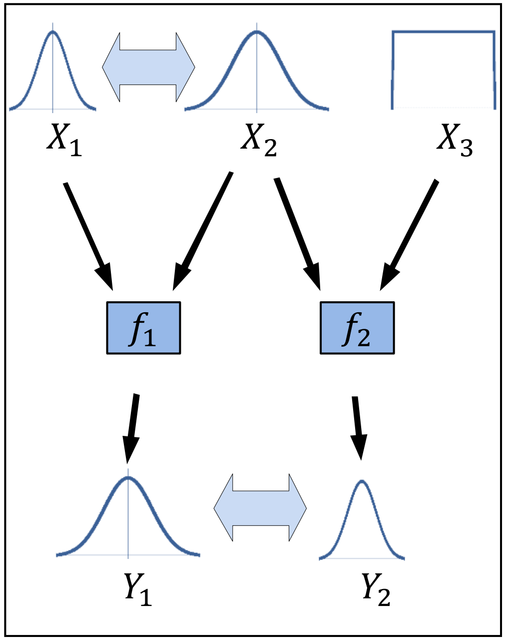
Ist die Korrelation positiv ($r(x_1, x_2) > 0$), nimmt die Unsicherheit ab.

Zusammenfassung

Allgemeines Modell mit:

- n Eingangsgrössen: $X_1, X_2, \ldots X_N$
- m Ausgangsgrössen: $Y_1, Y_2, \ldots Y_M$

$$\begin{aligned}
Y_1 &= f_1(X_1, X_2, \ldots, X_N) \\
Y_2 &= f_2(X_1, X_2, \ldots, X_N) \\
&\vdots \\
Y_M &= f_M(X_1, X_2, \ldots, X_N)
\end{aligned}$$



Verallgemeinerung auf das Gesetz der linearen Fortpflanzung von Unsicherheiten

$$u(y_k, y_l) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial f_k}{\partial x_i} \frac{\partial f_l}{\partial x_j} u(x_i, x_j)$$