Politechnika Wrocławska

Podstawy Metrologii i Technik Eksperymentu
Laboratorium

Analiza korelacyjna i regresyjna

Instrukcja do ćwiczenia nr 5

Zakład Miernictwa i Ochrony Atmosfery
Wrocław, kwiecień 2014

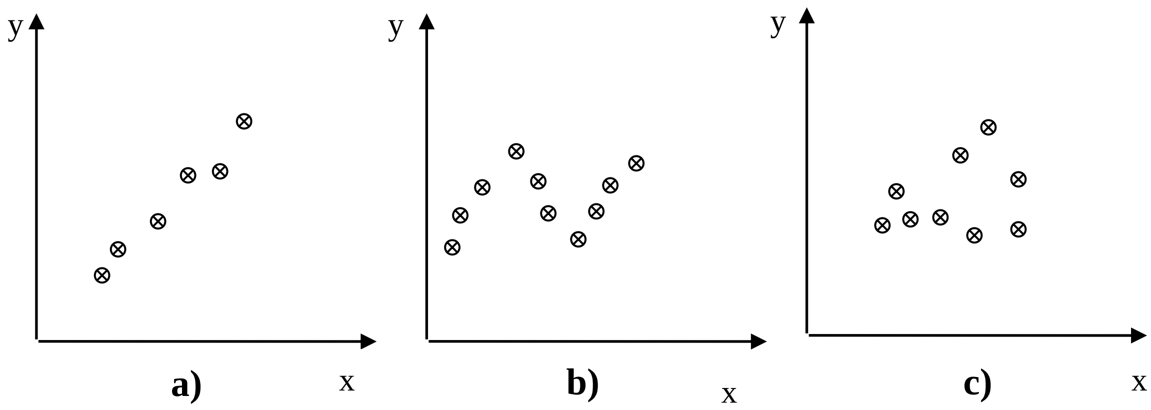
ANALIZA KORELACYJNA I REGRESYJNA

1. CEL ĆWICZENIA

Celem ćwiczenia jest obliczenie współczynnika korelacji serii pomiarów napięcia w funkcji temperatury zadanej dla przetworników temperatur z termoelementów typu K lub J oraz wyznaczenie charakterystyki tych przetworników za pomocą funkcji regresji. Dodatkowym celem jest sprawdzenie czy obliczone błędy pomiaru temperatury dla tych przetworników leżą poniżej błędów dopuszczalnych (granicznych) podanych przez producenta przetworników.

2. WSTĘP

Rysunek 1 przedstawia wg [1] wyniki pomiarów zmiennych $(x_1,y_1)\ldots(x_N,y_N)$ np. napięcia termoelektrycznego od temperatury, opopru od temperatury dla termometrów rezystancyjnych metalowych, zależność napięcia od natężenia przepływającego prądu czy strumienia objętości przepływającej cieczy od ciśnienia różnicowego na zwężce.



Rys. 1. Zależności między punktami pomiarowymi [1]: a) liniowe, b) krzywoliniowe, c) brak zależności

Widać, z niego że zależność między zmiennym na rysunku 1a jest liniowa, na rysunku 1b krzywoliniowa, a na rysunku 1c występuje brak zależności miedzy zmiennymi. Analiza korelacyjna pozwala określić zależności funkcyjne między zmiennymi.

W praktyce pomiarowej często ta zależność jest liniowa, a o liniowej zależności informuje **współczynnik korelacji liniowej** oznaczany przez **r** [2] i wyrażany tożsamymi równaniami:

$$r = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2 \sum(y_i-\bar{y})^2}} \tag{1}$$

$$r = \frac{\sum x_i y_i - N\bar{x}\,\bar{y}}{\sqrt{\left(\sum x_i^2 - N\bar{x}^2\right)\left(\sum y_i^2 - N\bar{y}^2\right)}} \tag{2}$$

Liczba r określa stopień zgodności punktów (x_i, y_i) z linią prostą i przyjmuje ona wartości z przedziału <-1,1>. Jeżeli r jest bliskie ±1 to punkty rozłożone są wzdłuż pewnej prostej, jeżeli r jest bliskie 0 to punkty są nieskorelowane i nie wyznaczają prostej [2].

Jeżeli współczynnik korelacji jest mniejszy niż 1 to z tabeli 1 [2] można obliczyć prawdopodobieństwo uzyskania współczynnika korelacji r większego od r_0 dla nieskorelowanych zmiennych x i y w zależności od liczby danych pomiarowych N, tzn. $P_N(|r| \geq r_0)$.

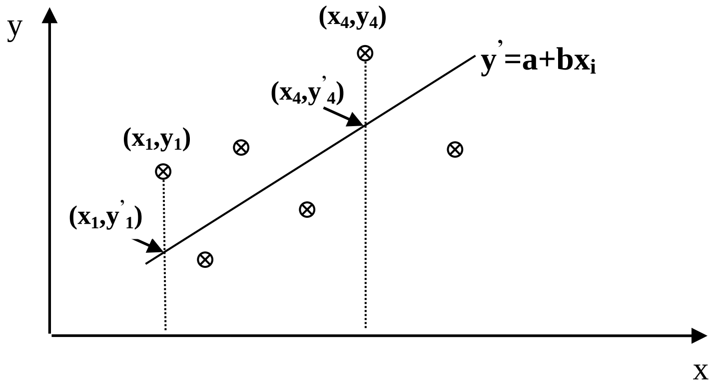
Tabela 1. Prawdopodobieństwo $P_N(|r| \geq r_0)$, że wyniki N pomiarów dwóch nieskorelowanych zmiennych x i y dały by współczynnik korelacji $|r| \geq r_0$. Podane wartości wyrażają prawdopodobieństwo procentowe, puste miejsca oznaczają wartości mniejsze niż 0,005 % [2].

	r_o										
N	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
3	100	94	87	81	74	67	59	51	41	29	0
6	100	85	70	56	43	31	21	12	6	1	0
10	100	78	58	40	25	14	7	2	0,5	–	0
20	100	67	40	20	8	2	0,5	0,1	–	–	0
50	100	49	16	3	0,4	–	–	–	–	–	0

Dla przykładu:

Dla 6 pomiarów (N=6) otrzymaliśmy r_0= 0,9; z tabeli 1 wynika, że jeżeli zmienne są nieskorelowane to prawdopodobieństwo uzyskania współczynnika korelacji większego od 0,9 wynosi 1%. Innymi słowy, jest bardzo mało prawdopodobne, że zmienne są nieskorelowane. Dla r_0= 0,5 prawdopodobieństwo, że zmienne są nieskorelowane wynosi już 31%.

Jeżeli uzyskana wartość współczynnika korelacji potwierdza liniową zależność między danymi pomiarowymi dwóch wielkości fizycznych x i y to można poprowadzić między punktami x_i i y_i prostą najlepiej do nich dopasowaną. Metoda analityczna znajdowania linii prostej, która najlepiej pasuje do szeregu punktów doświadczalnych nazywa się **metodą regresji liniowej lub metodą najmniejszych kwadratów**. Zasadę tę ilustruje rysunek 2



Rys. 2. Prosta regresji **y'=a+bx$_i$** między punktami pomiarowymi

Linię regresji prowadzi się tak, aby suma kwadratów różnic między (y_1 a y'_1) ….(y_4 a y'_4)…. (y_N a y'_N) była minimalna, tzn. szuka się minimum funkcji [1]:

$$e = \sum_{i=1}^{N}(y_i - y_i')^2 = \sum_{i=1}^{N}(y_i - a - bx_i)^2 = min \tag{3}$$

Wartości a i b otrzymane przez rozwiązanie układu równań $\frac{\partial e}{\partial a} = 0 \;\; i \;\; \frac{\partial e}{\partial b} = 0$ wynoszą:

$$b = \frac{\sum_{i=1}^{N} x_i y_i - \frac{1}{N}\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right)^2} \tag{5}$$

$$a = \frac{\sum_{i=1}^{N} y_i - b\sum_{i=1}^{n} x_i}{N} \tag{6}$$

Można wykazać, że średnie arytmetyczne [1]

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad i \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{7}$$

spełniają równanie: $\bar{y} = a + b\bar{x}$, wtedy stałe a i b wyrażają się równaniami:

$$b = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{8}$$

$$a = \bar{y} - b\bar{x} \tag{9}$$

Niepewności standardowe współczynników a i b oraz y' wyrażają się równaniami:

$$u_b = \sqrt{\frac{\sum_{i=1}^{N}(y_i - y')^2}{N-2}} \frac{1}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}} \tag{10}$$

$$u_a = \sqrt{\frac{\sum_{i=1}^{N}(y_i - y')^2}{N-2}} \sqrt{\frac{1}{N} + \frac{\bar{x}^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}} \tag{11}$$

3. SPOSÓB REALIZACJI ĆWICZENIA

Uwaga: na samym początku zajęć przypomnieć prowadzącemu o włączeniu urządzeń do prądu w celu ich rozgrzania. Czas rozgrzewania to ok 30 minut.

Badanym przetwornikiem temperatury jest przetwornik typu AR 580, przeznaczony dla termometru typu K lub J o następujących danych technicznych:

-zakres mierzonej temperatury 0-400°C

-zakres mierzonego napięcia 0- 10 V

-błąd graniczny Δg = 0,3% pełnego zakresu tj. Δg= 1,2°C

Z danych technicznych wynika, że charakterystykę przetwornika wyraża równanie:

$$\mathbf{U\,(V) = (10/400)\cdot t(^{\circ}C)} \qquad \mathbf{(11)}$$

$$\mathbf{t(^{\circ}C) = (400/10)\cdot U\,(V)} \qquad \mathbf{(12)}$$

Sposób realizacji ćwiczenia :

- dla nastawionych temperatur na kalibratorze C-402 t_i (i=1..9) równych odpowiednio: 400°C, 350°C, 300°C, 250°C, 250°C, 150°C ,100 °C, 50°C i 0 °C odczytać na multimetrze wartości napięcia $\mathbf{U_i}$ z przetwornika temperatury typu K lub typu J.
- z równania charakterystyki przetwornika (12) dla odczytanej wartości napięcia $\mathbf{U_i}$ (i=1..9) wyznaczyć wartość temperatury $\mathbf{t^*_i}$ (i=1...9)
- obliczyć odchyłkę δt_i= $\mathbf{t^*_i}$ **-** t_i (i=1..9)
- sprawdzić czy wartości odchyłki δt_i są mniejsze od błędu granicznego i narysować wykres $\delta t_i = f(t_i)$ - t_i równa się odpowiednio 400°C, 350°C, 300°C, 250°C, 250°C, 150°C ,100 °C, 50°C i 0 °C; na tym wykresie zaznaczyć również linię $\mathbf{\Delta g = \pm 1{,}2^{\circ}C}$

- dla zmierzonych serii (t_i, U_i) obliczyć z równania 1 lub 2 współczynnik korelacji **r.**

- dla zmierzonych serii (t_i, U_i) wyznaczyć charakterystykę przetwornika z funkcji regresji ; tzn. obliczyć współczynniki a i b z równań 5 i 6.

-podać następnie równanie charakterystyki przetwornika tj

$$\mathbf{U_R = a + b\cdot t} \qquad (13)$$

-z równań 10 i 11 obliczyć niepewności standardowe współczynników a i b

- równanie 13 przekształcić do postaci

$t = (U_R - a)/b,$ (14)

a następnie dla zmierzonych napięć Ui (i=1…9) obliczyć z równania 14 temperaturę $\mathbf{t_{Ri}}$

- obliczyć odchyłkę $\delta t_{Ri} = \mathbf{t_{Ri}} - t_i$ (i=1..9)

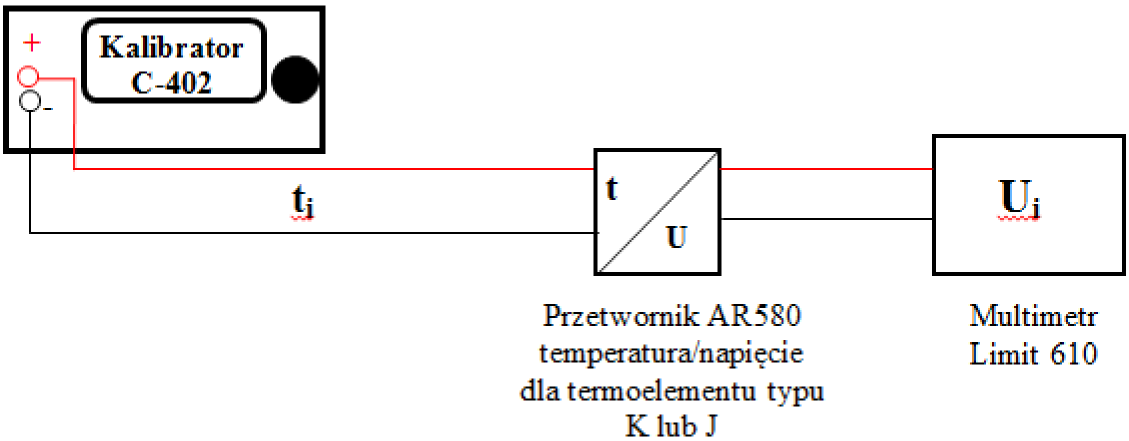
- sprawdzić czy wartości odchyłki δt_{Ri} są mniejsze od błędu granicznego i narysować wykres $\delta t_{Ri} = f(t_i)$ - t_i równa się odpowiednio 400°C, 350°C, 300°C, 250°C, 250°C, 150°C ,100 °C, 50°C i 0 °C; na tym wykresie zaznaczyć również linię **$\Delta g = \pm 1{,}2$°C.**

W trakcie realizacji ćwiczenia w kalibratorze powinny być wciśnięte przyciski:

- cal, t_o oraz J lub K w zależności od wybranego przetwornika

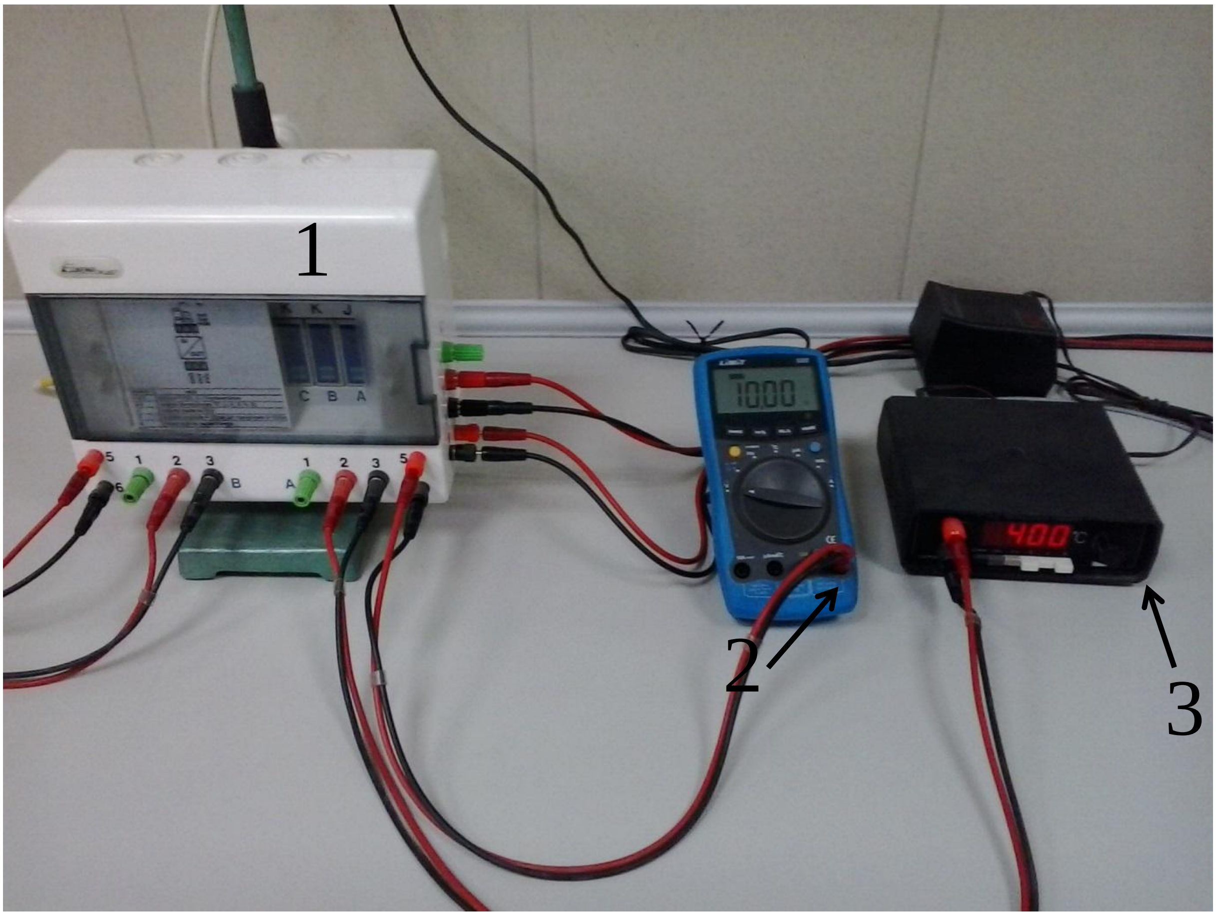
Rys. 3. Kalibrator C-402 oraz przetwornik AR580

Schemat stanowiska pomiarowego:



Rys. 4. Schemat stanowiska pomiarowego

Fotografia stanowiska pomiarowego:



Rys. 5. Fotografia stanowiska pomiarowego: 1- przetworniki temperatury dla termoelementu typu K i J, 2- multimetr Limit 610 (lub 600), 3- kalibrator C-402

Stanowisko pomiarowe składa się z:

Kalibrator C- 402:
błąd podstawowy ±0,1% wartości nastawionej ± 1°C
rozdzielczość ±1°C
zakres nastaw: typ J (Fe-CuNi): -210...+900°C
typ K (NiCr-NiAl): -150...+1370°C

Przetwornik AR 580:
zakres : 0 °C/0V - 400°C/10V
błąd podstawowy <0,3% pełnego zakresu pomiarowego
błąd rozdzielczości przetwarzania ±0,1°C× 100/zakres przetwarzania(°C)

Multimetr Limit 610
napięcie stałe: 60-600 mV-6-60-600-1000V
rozdzielczość: 0-10 mV
błąd graniczny: 1% wartości wskazanej + 3 cyfry

4. PYTANIA KONTROLNE

1. O czym informuje współczynnik korelacji liniowej?
2. Co oznacza, że r=±1 i r=0?
3. Ogólna zasada wyznaczania funkcji regresji
4. Definicja błędu granicznego
5. Napisać równanie charakterystyki przetwornika którego zakres odpowiada 0°C/0 V i 400 °C/10V

5. LITERATURA

1. Danuta Turzeniecka: *Ocena niepewności wyniku pomiaru,* Wydawnictwo Politechniki Poznańskiej 1977.
2. John R. Taylor: *Wstęp do analizy błędu pomiarowego*, PWN, Warszawa 1999.

Data wykonania instrukcji:
6.03.2014

Instrukcje wykonał: mgr inż. Michał Kamiński