Arbejdsseddel: Bevis for lineær regression

KBJ, januar 2025 3u MA

**Udangspunkt**

Vi får givet et datasæt med datapunkter:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Vi søger en lineær funktion som model over vores data på formen:

Vi vil i det følgende bruge sumtegn i betydningen:

Bemærk at vi forkorter sum-tegns-notationen, så vi beslutter at det altid er summen fra 1 til uden at skrive dette eksplicit. Ellers bliver det følgende ulideligt.

**Opgave 1**

Vi betegner middelværdien af alle -værdier i tabellen (udtales -streg) og tilsvarende middelværdien af alle y-værdier med .

1. Opskriv udtryk for og ved sumtegn, og udtryk derpå summerne og ved deres middelværdier.

**Opgave 2**

1. Mind dig selv om hvad begrebet *residual* dækker over.
2. Opstil et udtryk for residual nr. , dvs dels ved dels ved og .
3. Bestem et udtryk for det kvadrerede residual - ophæv alle parenteser.

**Opgave 3**

Den bedste rette linje er den linje, hvor summen af kvadrerede residualer bliver mindst mulig.

Vi indfører derfor funktionen ”Sum of Squared Errors”, kaldet :

1. Hvorfor er en funktion af og ?
2. Opskriv den fulde forskrift for .
3. Hvorfor er vi særligt interesserede i *minimum* for ?

**Opgave 4**

Vi ønsker nu at finde *minimum* for :

1. Bestem partialt afledede og .
2. Bestem stationære punkter for ved at løse ligningssystemet og . (Hint! Start med anden ligning).
3. Hvilke antagelser må der laves undervejs.
4. Hvor mange stationære punkter har ?

**Opgave 5**

1. Undersøg at det stationære punkt for faktisk er af arten ”et minimum”.

**Opgave 6**

1. Udfør manuel lineær regression på følgende data, med formlerne udledt i opgave 4.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

1. Tjek med Nspire at det passer.