Huiswerk Micro, serie 1, 8-2-12

Inleveren: Deze week is er geen inleveropgave.

Opgave 1. Beschouw het lineaire model $y \approx \theta x + \theta'$. Stel dat hiervoor N waarnemingen $(x^{(i)}, y^{(i)})$ zijn verricht aan het paar (x, y). Laat zien dat de formules voor $\hat{\theta}$ en $\hat{\theta}'$ op pp. 37-38 van de syllabus ook worden bereikt volgens de methode van Example 1.1.1 (partiële afgeleiden van de kwadratensom $SS(\theta, \theta')$ gelijkstellen aan nul).

Opgave 2. Beschouw hetzelfde model als in opgave 1. Stel dat het laatste paar waarnemingen $(x^{(N)}, y^{(N)})$ "gigantisch groot" is ten opzichte van de overige $N-1$ paren. Wat kan dan gezegd worden van $\hat{\theta}$ en $\hat{\theta}'$? N.B. Het totale aantal waarnemingen verandert niet! Analyseer dit vraagstuk wiskundig door "gigantisch groot" alsvolgt te formaliseren.
Geval (i): $x^{(N)} \to \infty$, $y^{(N)} \to \infty$ en $y^{(N)}/x^{(N)} \to \alpha \in \mathbb{R}_+$,
Geval (ii) $x^{(N)} \to \infty$, $y^{(N)} \to \infty$ en $y^{(N)}/x^{(N)} \to +\infty$.

Hint: Ongeacht de grootte van $x^{(N)}$ en $y^{(N)}$ blijft gelden dat de regressielijn door het punt $(m_{\mathbf{x}}, m_{\mathbf{y}})$ moet gaan: zie p. 38 van de syllabus. *Attentie:* op p. 38 van de uitgedeelde syllabus staat een drukfout: zie de verbeteringen via de thuispagina.

Opgave 3. Beschouw het algemene lineaire model uit sectie 2.2. Stel dat de design matrix X weer de vorm $X = (\mathbf{x}^1 \mid \cdots \mid \mathbf{x}^d)$ heeft, maar dat nu extra geldt $\mathbf{x}^d = \mathbf{1}$, net als in Example 2.2.6. Toon aan dat in dit geval de som $\sum_{i=1}^{N} \hat{e}_i$ gelijk is aan nul. Waar is deze eigenschap terug te vinden in Example 2.2.6?

Opgave 4. Accountants moeten vaak de door hen gevonden waarden van te controleren bedrijfsgoederen, zoals meubilair, wagenpark, etc., vergelijken met hun respectieve boekwaarden, d.w.z. de waarden waarvoor de firma ze in haar boeken heeft staan. Als een firma goed boekhoudt, zal er sprake zijn van een sterk lineair verband tussen beide. Een bepaalde firma heeft in zijn boeken de volgende waarden $x^{(i)}$ voor tien goederen staan: 10, 12, 9, 27, 47, 112, 36, 241, 59 en 167. Bij controle vinden de accountants de volgende waarden $y^{(i)}$: 9, 14, 7, 29, 45, 109, 40, 238, 60, 170. *Hint 1:* Hier is $\sum_{i=1}^{10}(x^{(i)})^2 = 106554$ en $\sum_{i=1}^{10} x^{(i)} y^{(i)} = 106155$.
a. Welke toename in de gecontroleerde waarde verwacht je van een "gemiddeld" bedrijfsgoed bij een toename van 1 €van de boekwaarde?
b. Stel een goed heeft een boekwaarde van 100 €. Schat de gevonden waarde ervan op basis van de kleinste kwadratenmethode. Gebruik zowel het model waarbij de regressielijn door de oorsprong gaat (dus met $d = 1$) als het model waarbij dat niet is vereist ($d = 2$). Doe dit eerst zonder gebruik te maken van de computer, en controleer

vervolgens je uitkomsten met een programma zoals Leastsquares (zie voetnoot 2 in de syllabus).

Hint 2: Ter controle: de regressielijn voor het geval $d = 2$ blijkt $y = 0.9914x + 0.7198$ te zijn.

Opgave 5. Stel dat de design matrix voor het lineaire model (2.5) volle rang heeft, zodat formule (2.11) opgaat. Toon nu *rechtstreeks* aan dat ongelijkheid (2.9) geldt, en wel door $\mathbf{c} := \mathbf{y} - X\boldsymbol{\theta}$ te onbinden als som van $\mathbf{a} := \mathbf{y} - X\hat{\boldsymbol{\theta}}$ en $\mathbf{b} := X\hat{\boldsymbol{\theta}} - X\boldsymbol{\theta}$ en vervolgens *alleen* met behulp van matrix-vectorrekening aan te tonen dat $\|\mathbf{a}\|^2 \leq \|\mathbf{c}\|^2$.

Opgave 6. Zoals bekend van het college Infinitesimaalrekening, is een noodzakelijke voorwaarde voor het feit dat een differentieerbare functie $f(\boldsymbol{\theta})$ over $\mathbb{R}^d$ zijn minimum aanneemt een punt $\hat{\boldsymbol{\theta}}$ dat de gradient $\nabla f(\boldsymbol{\theta})$ gelijk is aan nul. Bepaal de gradient van $f(\boldsymbol{\theta}) := \|\mathbf{y} - X\boldsymbol{\theta}\|^2$ en bewijs vervolgens dat bovenstaande noodzakelijke voorwaarde equivalent is met het oplossen van het stelsel normaalvergelijkingen $(\mathbf{y} - X\boldsymbol{\theta}) \cdot \mathbf{x}^j = 0$ voor alle $j = 1, \ldots, d$.