Dr Wesseling

INSTITUUT VOOR CULTUURTECHNIEK EN WATERHUISHOUDING
NOTA No.142 d.d.5 september 1962

Een aantekening bij het bepalen van afgeleiden van functies afkomstig uit matrixprodukten

Ph.Th.Stol

In de vectoralgebra en matrixrekening komen produktvormen voor die een scalaire grootheid of een functie voorstellen.

Zo kan, omgekeerd, de functie

$$y = a_1 x_1 + a_2 x_2 + a_3 x_3 \tag{1}$$

opgevat worden als het inproduct van de kolom vectoren

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{en} \quad \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

De functie wordt dan geschreven in de volgende vorm (waarvoor de commutatieve eigenschap geldt):

$$y = \underline{a} \cdot \underline{x} = \underline{x} \cdot \underline{a} \tag{2}$$

of in matrixnotatie, met t als aanduiding voor het transponeren

$$y = {}^t\underline{a}\,\underline{x} = {}^t\underline{x}\,\underline{a}$$

In het bovenstaande representeert (1) en daarmee (2) een functie. Gevraagd kan worden van deze functie de achtereenvolgende afgeleiden naar x_i te bepalen en deze in een kolom te rangschikken. De bewerking wordt symbolisch voorgesteld door

$$\frac{d y}{d \underline{x}} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial x_n} \end{pmatrix} y$$

waarin de vector een operator is. Voor het driedimensionale geval komt deze overeen met de operator nabla (∇) uit de vectoranalyse en het antwoord op de gestelde vraag is dan de gradiënt van de functie en wel, als gegeven is

$$\varphi = \varphi\ (x_1, x_2, x_3) = \text{constant}$$

de vector

$$\nabla\varphi = \begin{pmatrix} \frac{\partial \varphi}{\partial x_1} \\ \frac{\partial \varphi}{\partial x_2} \\ \frac{\partial \varphi}{\partial x_3} \end{pmatrix} \tag{3}$$

17870 66

Wordt φ verkregen uit een vector- of matrixprodukt dan ontstaat de vraag (3) in de oorspronkelijke elementen (vector, matrix) uit te drukken. De regels waarmee de gevraagde uitdrukkingen worden verkregen, worden hierna voor enkele gevallen in het kort aangegeven.

Het inprodukt

1) $$y = {}^t\underline{a}\,\underline{x}$$
$$= a_1x_1 + a_2x_2 + a_3x_3$$
$$\nabla y = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \underline{a}$$

zodat

$$\frac{dy}{d\underline{x}} = {}^t(\,{}^t\underline{a}\,) = \underline{a}$$

2) $$y = {}^t\underline{x}\,\underline{a}$$

Doordat de commutatieve eigenschap voor inprodukten geldt volgt direct

$$\frac{dy}{d\underline{x}} = \underline{a}$$

De lengte-kwadraat

3) Uit 1) en 2) volgt nu voor

$$y = \underline{x}^2 = \underline{x}\cdot\underline{x} = {}^t\underline{x}\,\underline{x}$$

dat

$$\frac{dy}{d\underline{x}} = \underline{x} + {}^t({}^t\underline{x}) = 2\,\underline{x}$$

Sommen van inprodukten

4) $$y = \underline{a}\cdot\underline{x} + \underline{b}\cdot\underline{x}$$
$$= (\underline{a}+\underline{b})\cdot\underline{x}$$
$$= {}^t(\underline{a}+\underline{b})\,\underline{x}$$

zodat

$$\frac{dy}{d\underline{x}} = (\underline{a}+\underline{b})$$

5) $$V = \underline{a}\cdot\underline{x} + \underline{b}\cdot\underline{y}$$
$$= {}^t\underline{a}\underline{x} + {}^t\underline{b}\underline{y}$$

waarvoor geschreven kan worden

$$V = \begin{pmatrix} {}^t\underline{a} & {}^t\underline{b} \end{pmatrix}\begin{pmatrix} \underline{x} \\ \underline{y} \end{pmatrix}$$

De operator heeft nu de gedaante

$$\begin{pmatrix} \frac{\partial}{\partial x_1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial x_n} \\ \frac{\partial}{\partial y_1} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial y_n} \end{pmatrix} \underset{===}{\overset{def}{}} \frac{d}{d\underline{r}}$$

zodat, volgens 1) geldt

$$\frac{d\underline{y}}{d\underline{r}} = {}^t({}^t\underline{a} \quad {}^t\underline{b}) = \begin{pmatrix} \underline{a} \\ \underline{b} \end{pmatrix}$$

Kwadratische vormen

6) $$y = (\underline{a} \cdot \underline{x})^2 = ({}^t\underline{a}\,\underline{x})^2 = (a_1x_1 + a_2x_2 + a_3x_3)^2$$

Uit de beide uiterste leden volgt

$$\frac{d\underline{y}}{d\underline{x}} = \begin{pmatrix} 2(\underline{a} \cdot \underline{x})\, a_1 \\ 2(\underline{a} \cdot \underline{x})\, a_2 \\ 2(\underline{a} \cdot \underline{x})\, a_3 \end{pmatrix} = 2(\underline{a} \cdot \underline{x})\,\underline{a}$$

Deze uitkomst volgt ook uit 1) door toepassen van de produktregel.

Opgemerkt wordt dat de produktregel slechts op het associatieve matrixprodukt toegepast mag worden dus

$$y = {}^t({}^t\underline{a}\,\underline{x})({}^t\underline{a}\,\underline{x})$$
$$= {}^t\underline{x}\,\underline{a}\,{}^t\underline{a}\,\underline{x}$$

en

$$\frac{d\underline{y}}{d\underline{x}} = \underline{a}\,{}^t\underline{a}\,\underline{x} + {}^t({}^t\underline{x}\,\underline{a}\,{}^t\underline{a})$$
$$= 2\,\underline{a}\,{}^t\underline{a}\,\underline{x}$$

Herleid op de uitgangsvorm wordt dit

$$\frac{d\underline{y}}{d\underline{x}} = 2\,\underline{a}\,(\underline{a} \cdot \underline{x})$$

Tweede graadsvormen met matrices

7) De algemene gedaante van een tweede graadsvorm is in matrixnotatie (voor een tweedimensionaal systeem)

$$y = {}^t\underline{x}\ A\ \underline{x}$$

$$= (x_1, x_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= (x_1, x_2)\begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}$$

$$= ax_1^2 + (b + c)\ x_1x_2 + dx_2^2$$

Hieruit volgt dat

$$\frac{d\underline{y}}{d\underline{x}} = \begin{pmatrix} 2ax_1 + (b + c)\ x_2 \\ (b + c)\ x_1 + 2\ dx_2 \end{pmatrix}$$

$$= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\underline{x} + \begin{pmatrix} a & c \\ b & d \end{pmatrix}\underline{x}$$

zodat

$$\frac{d\underline{y}}{d\underline{x}} = A\,\underline{x} + {}^tA\,\underline{x}$$

Door toepassing van de produktregel op y wordt deze uitkomst rechtstreeks verkregen namelijk

$$\frac{d}{d\underline{x}}\ {}^t\underline{x}\ A\ \underline{x} = A\,\underline{x} + {}^t(\,{}^t\underline{x}\,A)$$

$$= A\,\underline{x} + {}^tA\,\underline{x}$$

8) Voor het geval dat matrix A uit 7) symmetrisch is geldt $A = {}^tA$ zodat in dat geval de uitkomst luidt

$$\frac{d}{d\underline{x}}\ {}^t\underline{x}\,A\,\underline{x} = 2\,A\,\underline{x}$$

<u>Kwadratische vormen met matrices</u>

9)

$$y = (\,A\,\underline{x}\,)^2$$

$$= \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]^2 \qquad (4)$$

$$= \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}^2$$

$$= (ax_1 + bx_2)^2 + (cx_1 + dx_2)^2$$

Hieruit volgt voor de afgeleide

$$\frac{dy}{d\underline{x}} = \begin{pmatrix} 2\,(ax_1 + bx_2)\,a + 2\,(cx_1 + dx_2)\,c \\ 2\,(ax_1 + bx_2)\,b + 2\,(cx_1 + dx_2)\,d \end{pmatrix}$$

$$= 2 \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix}$$

$$= 2\,{}^{t}A\,A\,\underline{x}$$

De uitkomst kan ook met de produktregel verkregen worden door (4) als matrixprodukt te schrijven en 8) toe te passen. De afleiding wordt dan

$$y = (A\,\underline{x})^2 = {}^{t}(A\,\underline{x})\,(A\,\underline{x})$$

$$= {}^{t}\underline{x}\;{}^{t}A\,A\,\underline{x}$$

stel $\quad {}^{t}A\,A = B$

waarin B een symmetrische matrix is dan komt er

$$\frac{d}{d\underline{x}}\;{}^{t}\underline{x}\,B\,\underline{x} = 2\,B\,\underline{x}$$

$$= 2\,{}^{t}A\,A\,\underline{x}$$

In het voorgaande werden enkele regels gegeven voor het bepalen van afgeleiden van scalaire functies die uit vector- respectievelijk matrixprodukten zijn ontstaan. De afgeleiden kunnen bepaald worden door naar de variabele vector te differentiëren waarbij bedacht moet worden:

1. dat de produktregel slechts kan worden toegepast wanneer de functie als matrixprodukt geschreven wordt (waarvoor de associatieve eigenschap geldt)
2. dat vormen vóór de variabele in hun getransponeerden overgaan, en vormen achter de variabele in hun oorspronkelijke vorm in de afgeleide voorkomen.
3. dat de kettingregel niet altijd toegepast kan worden (vergelijk 6) met 9))

De meetkundige betekenis van de uitkomsten is deze dat de richting van de normaal op het oppervlak φ = constant verkregen wordt. (SPIEGEL, pag. 62 e.v., TIMMAN, par. 67 e.v.). Wordt deze normaal voorgesteld door de $\underline{n}$ dan is dus

$$\underline{n} = \nabla\,\varphi \quad \text{of} \quad \underline{n} = \frac{d\,\varphi}{d\,\underline{x}}$$

De mate van verandering van de functiewaarde naar hogere waarden wordt gegeven door de lengte van de normaal namelijk

$$|\underline{n}| = |\nabla\varphi|$$

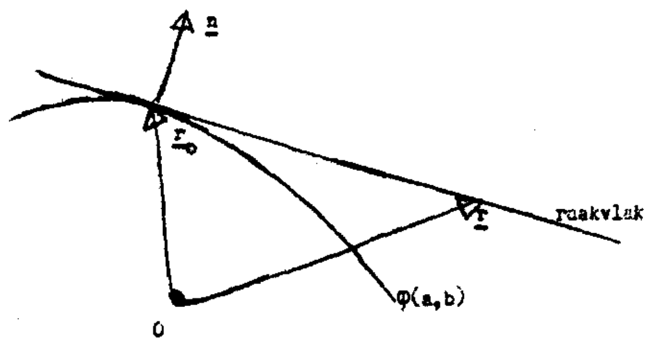
Wordt φ voorgesteld door niveaulijnen in het grondvlak dan zullen bij grote $\underline{n}$ de niveaulijnen dicht bijeen liggen. Is $\underline{n}$ klein, dan liggen deze ver uiteen.

Is $\underline{r}_o$ het punt van φ waarvoor $\underline{n}$ berekend wordt dan geldt voor het raakvlak in dat punt aan φ

$$(\underline{r} - \underline{r}_o) \cdot \underline{n} = 0$$

waarin de vector $\underline{r} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$ de coördinaten van een willekeurig punt van het raakvlak representeert.

In schema



Ten slotte kan nog gevraagd worden naar de mate van verandering van de functiewaarde φ in een gegeven richting bijvoorbeeld $\underline{a}$. De zogenaamde richtings-afgeleide is in dit geval het inprodukt

$$\nabla \varphi \cdot \frac{\underline{a}}{|\underline{a}|}$$

of, met $\underline{a}_e$ als aanduiding van een eenheidsvector in de richting van $\underline{a}$ waarvoor $|\underline{a}_e| = 1$,

$$\nabla \varphi \cdot \frac{\underline{a}}{|\underline{a}|} = \underline{n} \cdot \underline{a}_e$$

Meetkundig stelt deze vorm de component van $\underline{n}$ in de richting van $\underline{a}$ voor daar voor het inprodukt geldt

$$\underline{n} \cdot \underline{a}_e = |\underline{n}|\,|\underline{a}| \cos \alpha = |n| \cos \alpha$$

welke laatste vorm de projectie van de vector $\underline{n}$ op de vector $\underline{a}$ voorstelt.

In het bijzonder geldt nog voor de raaklijn aan een niveaulijn en dus met $\underline{a}_e$ loodrecht op $\underline{n}$ dat

$$\underline{n} \cdot \underline{a}_e = 0$$

Er treedt dus in deze richting (over een klein afstandje) geen verandering op.

Literatuur SPIEGEL, M.R. (1959) Theory and problems of vector analysis and an introduction to tensor analysis. New York (I.C.W. 11/124)

TIEMAN, R (1959) Vectoranalyse. Delft (I.C.W. 11/93)