10. Calcolo Differenziale, II

10.a Differenziale e matrice jacobiana

10.1 Definizione. *Sia* $A \subset \mathbb{R}^n$, $n \geq 1$, *aperto e* $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, $f = (f^1, f^2, \dots, f^m)$, $m \geq 1$. f *si dice* differenziabile *in* $x_0 \in A$ *se esiste una applicazione lineare* $L : \mathbb{R}^n \to \mathbb{R}^m$, $L = (L^1, L^2, \dots, L^m)$, *detta* differenziale di f in x_0 *o* applicazione lineare tangente ad f in x_0, *tale che*

$$\frac{f(x_0+h) - f(x_0) - L(h)}{|h|} \to 0 \qquad \text{per } h \to 0. \tag{10.1}$$

Se f *è differenziabile in ogni punto di* A, *si dice che* f *è* differenziabile in A.

Poichè la (10.1) è il sistema di m limiti

$$\frac{f^i(x_0+h) - f^i(x_0) - L^i(h)}{|h|} \to 0 \qquad \text{per } h \to 0, \tag{10.2}$$

per ogni $i = 1, \dots, m$, si conclude che $f = (f^1, f^2, \dots, f^m)$ è differenziabile in x_0 con differenziale $L = (L^1, L^2, \dots, L^m)$ se e solo se per ogni $i = 1, \dots, m$ $f^i(x)$ è differenziabile in x_0 con differenziale L^i e dunque

$$L^i(h) = \frac{\partial f^i}{\partial h}(x_0) = \sum_{j=1}^n \frac{\partial f^i}{\partial x^j}(x_0) h^j \qquad \forall i = 1, \dots, m. \tag{10.3}$$

Introducendo la matrice $m \times n$, detta *matrice jacobiana di* f *in* x_0,

$$\mathbf{D}f(x_0) := \Big[\frac{\partial f^i}{\partial x^j}(x_0)\Big] = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \frac{\partial f^1}{\partial x^2}(x_0) & \dots & \frac{\partial f^1}{\partial x^n}(x_0) \\ \frac{\partial f^2}{\partial x^1}(x_0) & \frac{\partial f^2}{\partial x^2}(x_0) & \dots & \frac{\partial f^2}{\partial x^n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x^1}(x_0) & \frac{\partial f^m}{\partial x^2}(x_0) & \dots & \frac{\partial f^m}{\partial x^n}(x_0) \end{pmatrix},$$

si ha

$$L(h) = \begin{pmatrix} L^1(h) \\ L^2(h) \\ \vdots \\ L^m(h) \end{pmatrix} = \mathbf{D}f(x_0)h.$$

In conclusione,

10.2 Proposizione. *Sia $A \subset \mathbb{R}^n$ aperto e $f : A \to \mathbb{R}^m$ una funzione, $f = (f^1, f^2, \dots, f^m)$. f è differenziabile in $x_0 \in A$ se e solo se*

(i) *le componenti f^i, $i = 1, \dots, m$, hanno tutte le derivate parziali in x_0,*
(ii) *detta $\mathbf{D}f(x_0)$ la matrice jacobiana in* (9.5), *si ha*

$$f(x_0 + h) - f(x_0) - \mathbf{D}f(x_0)h = o(|h|) \qquad \text{per } h \to 0.$$

Il differenziale di $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ in x_0 si indica anche con $df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ e dunque

$$df(x_0)(h) := \mathbf{D}f(x_0)h \qquad h \in \mathbb{R}^n. \tag{10.4}$$

10.b Spazio tangente al grafico

Sia $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ differenziabile in un punto x_0 interno ad A. Identifichiamo $\mathbb{R}^n \times \mathbb{R}^m$ con $\mathbb{R}^{n+m}$. Il *grafico* di f è definito come il sottoinsieme di $\mathbb{R}^n \times \mathbb{R}^n$ dato da

$$G_f := \Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, y = f(x)\Big\}$$

Chiamiamo *piano tangente al grafico di f in x_0* il sottoinsieme grafico dell'applicazione

$$\Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,\Big|\, y = f(x_0) + \mathbf{D}f(x_0)(x - x_0)\Big\}$$

e *piano tangente al grafico di f in* $(x_0, f(x_0))$ il suo traslato nell'origine di $\mathbb{R}^n \times \mathbb{R}^m$.

$$\mathrm{Tan}_{(x_0, f(x_0))} G_f := \Big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \,\Big|\, y = \mathbf{D}f(x_0)x\Big\},$$

Ovviamente lo spazio tangente al grafico in $(x_0, f(x_0))$ è anche il grafico dell'applicazione lineare tangente ad f in x_0

$$x \to L(x) = \mathbf{D}f(x_0)x.$$

Chiaramente $\mathrm{Tan}_{(x_0, f(x_0))}$ è l'immagine della mappa lineare iniettiva da $\mathbb{R}^n$ in $\mathbb{R}^n \times \mathbb{R}^m$ data da $x \to (x, \mathbf{D}f(x_0)x)$, i.e., in termini di matrici, $\mathrm{Tan}_{(x_0, f(x_0))} G_f = \mathrm{Im}\,\mathbf{A}$ dove si è posto

$$\mathbf{A} := \left(\begin{array}{|c|} \hline \\ \mathrm{Id} \\ \\ \hline \mathbf{D}f(x_0) \\ \hline \end{array}\right).$$

Evidentemente le colonne di $\mathbf{A}$ sono linearmente indipendenti; sono quindi una base di $\mathrm{Tan}_{(x_0, f(x_0))} G_f$.

10.c Spazio normale al grafico

$\mathrm{Tan}_{(x_0,f(x_0))}$ è caratterizzato come l'insieme dei vettori di $\mathbb{R}^n \times \mathbb{R}^m$ tali che $\mathbf{D}f(x_0) - y = 0$. Introducendo la matrice $m \times (n+m)$

$$\mathbf{B} := \left(\boxed{\quad \mathbf{D}f(x_0) \quad} \quad \boxed{\quad -\,\mathrm{Id} \quad}, \right) \tag{10.5}$$

l'equazione $\mathbf{D}f(x_0) - y = 0$ si riscrive come

$$\mathbf{B}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

e quindi

$$\mathrm{Tan}_{(x_0,f(x_0))} G_f = \ker \mathbf{B}.$$

Se usiamo il prodotto scalare standard in $\mathbb{R}^{n+m}$ e ci ricordiamo che abbiamo identificato $\mathbb{R}^n \times \mathbb{R}^m$ con $\mathbb{R}^{n+m}$, si conclude che

10.3 Proposizione. *Le m righe della matrice* (10.5), *pensati come vettori di $\mathbb{R}^{n+m}$, sono una base dello spazio ortogonale a* $\mathrm{Tan}_{(x_0,f(x_0))} G_f$.

10.d Qualche exempio

10.4 Esercizio (Curve in $\mathbb{R}^m$). Le mappe $r : [a,b] \subset \mathbb{R} \to \mathbb{R}^m$, $m \geq 1$, si chiamano anche *curve* di $\mathbb{R}^m$, il senso è quello di una *parametrizzazione* dell'immagine o *traiettoria* della curva. Se $r(t) = (r^1(t), \dots, r^m(t))$, $r(t)$ è differenziabile se e solo se tutte le sue componenti sono derivabili. Calcolare la matrice jacobiana e l'applicazione lineare tangente, calcolare la curva grafico di r, il suo spazio tangente e quello normale in un punto.
Soluzione. La matrice jacobiana di r è la matrice di dimensione $m \times 1$

$$\mathbf{D}r(t) = \begin{pmatrix} {r^1}'(t) \\ \vdots \\ {r^m}'(t) \end{pmatrix} =: r'(t).$$

$\mathbf{D}r(t)$ è detta la *velocità* della curva al tempo t.

L'applicazione lineare tangente in t_0 è la mappa $t \to r'(t_0)t$. Percio' l'applicazione lineare tangente è l'equazione parametrica della retta tangente alla curva $r(t)$ in $r(t_0)$ nel caso in cui r è iniettiva e $r'(t_0) \neq 0$.

Il grafico di $r(t)$ è la curva di $\mathbb{R} \times \mathbb{R}^m$

$$t \to \begin{pmatrix} t \\ r(t) \end{pmatrix}$$

e lo spazio tangente al grafico di r in t_0 è la retta di $\mathbb{R} \times \mathbb{R}^m$ per l'origine data dall'immagine di

$$t \to \begin{pmatrix} t \\ r'(t_0)t \end{pmatrix} = \begin{pmatrix} 1 \\ r'(t_0) \end{pmatrix} t.$$

Il piano normale al grafico è generato dagli m vettori riga della matrice

$$\begin{pmatrix} {r^1}'(t) & -1 & 0 & \dots & 0 \\ {r^2}'(t) & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ {r^m}'(t) & 0 & 0 & \dots & -1. \end{pmatrix}$$

10.5 Esercizio (Superfici parametrizzate). Siano $A \subset \mathbb{R}^2$ e (u, v) le coordinate in $\mathbb{R}^2$. Mappe $r : A \to \mathbb{R}^3$, $r(u, v) = (x(u, v), y(u, v), z(u, v))^T$, sono utili a descrivere "superfici" in $\mathbb{R}^3$, guardando ad r come ad una *parametrizzazione* della sua immagine $r(A) \subset \mathbb{R}^m$. Descrivere matrice jacobiana, grafico, spazi tangente e normale al grafico di una superficie parametrizzata.
Soluzione. Siano $A \subset \mathbb{R}^2$, (u, v) le coordinate in $\mathbb{R}^2$ e $r : A \to \mathbb{R}^3$, $r(u, v) = (x(u, v), y(u, v), z(u, v))^T$. Se r è differenziabile in (u_0, v_0), la sua matrice jacobiana è

$$\mathbf{D}r(u_0, v_0) = \begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix}.$$

dove si sono abbreviate le derivate parziali come

$$x_u := \frac{\partial x}{\partial u}(u_0, v_0), \quad x_v := \frac{\partial x}{\partial v}(u_0, v_0), \quad \dots.$$

I vettori colonna della matrice jacobiana, $r_u := (x_u, y_u, z_u)^T$, $r_v := (x_v, y_v, z_v)^T$ sono rispettivamente i vettori velocità delle curve $u \to r(u, v_0)$ e $v \to r(u_0, v)$, rispettivamente in $u = u_0$ e $v = v_0$. Le curve $u \to r(u, v_0)$ e $v \to r(u_0, v)$ sono a loro volta immagine tramite r delle rette $u \to (u, v_0)$ e $v \to (u_0, v)$ di $\mathbb{R}^2$.

Lo spazio tangente al grafico di $r(u, v)$ in (u_0, v_0) è il sottospazio due dimensionale di $\mathbb{R}^2 \times \mathbb{R}^3$ grafico della mappa

$$\begin{pmatrix} u \\ v \end{pmatrix} \to \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Una sua base in $\mathbb{R}^5$ è data dalle due colonne della matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix}.$$

Infine le tre righe della matrice

$$\begin{pmatrix} x_u & x_v & -1 & 0 & 0 \\ y_u & y_v & 0 & -1 & 0 \\ z_u & z_v & 0 & 0 & -1 \end{pmatrix}$$

sono una base del piano normale allo spazio tangente.

10.6 Esercizio. Calcolare l'applicazione lineare tangente a $x \to \mathbf{A}x$ dove $\mathbf{A} = [a^i_j] \in M_{m,n}(\mathbb{R})$.
Soluzione. Sia $f(x) = \mathbf{A}x$, i.e., $f = (f^1, f^2, \dots, f^m)$, $f^i(x) = \sum_{j=1}^n a^i_j x^j$. Per ogni $h = 1, \dots, n$

$$\frac{\partial f^i}{\partial x^h}(x) = \sum_{j=1}^n a^i_j \delta^j_h = a^i_h$$

e dunque con notazione matriciale

$$\mathbf{D}(\mathbf{A}x) = \mathbf{A} \qquad \forall x \in \mathbb{R}^n.$$

10.7 Esercizio. Calcolare la matrice jacobiana della forma quadratica $\phi(x) = (\mathbf{A}x) \bullet x = x^T \mathbf{A}x$ dove $\mathbf{A} = [a_{ij}] \in M_{n,n}(\mathbb{R})$.
Soluzione. In coordinate,

$$\phi(x) = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

Per ogni $h = 1, \dots, n$

$$\frac{\partial \phi}{\partial x_h}(x) = \sum_{i,j=1}^{n} a_{ij}(\delta_{ih} x_j + x_i \delta_{jh}) == \sum_{j=1}^{n} a_{hj} x_j + \sum_{i=1}^{n} a_{ih} x_i = (\mathbf{A}x)_h + (\mathbf{A}^T x)_h$$

e dunque

$$\mathbf{D}\phi(x) = (\mathbf{A} + \mathbf{A}^T)x.$$

Se $\mathbf{A}$ è simmetrica, $\mathbf{A} = \mathbf{A}^T$, allora $\mathbf{D}\phi(x) = 2\mathbf{A}x$.

10.8 Campi di vettori. L'assegnazione in ogni punto x di un aperto A di un vettore $v(x) \in \mathbb{R}^m$, cioè di un *campo di vettori*, si descrive evidentemente come una mappa $f : A \to \mathbb{R}^m$ $f = (f^1, f^2, \dots, f^m)$. Si pensi ad esempio al campo delle velocità delle particelle di un fluido, al campo elettrostatico, o al campo gravitazionale. Due *operatori differenziali* che hanno particolare rilevanza nel contesto dei campi vettoriali sono l'operatore di *divergenza*,

$$\operatorname{div} f = \nabla \bullet f := \sum_{i=1}^{n} \frac{\partial f^i}{\partial x^i},$$

e, per $n = 3$, l'operatore di *rotore*

$$\operatorname{rot} f = \nabla \times f := \left(\frac{\partial f^3}{\partial x^2} - \frac{\partial f^2}{\partial x^3}, \frac{\partial f^1}{\partial x^3} - \frac{\partial f^3}{\partial x^1}, \frac{\partial f^2}{\partial x^1} - \frac{\partial f^1}{\partial x^2} \right).$$

Un campo vettoriale $f : A \subset \mathbb{R}^n \to \mathbb{R}^n$ si dice *solenoidale* se $\operatorname{div} f = 0$, e *irrotazionale*, se $\operatorname{rot} f = 0$.

10.9 Proposizione. *Se $f, g : A \to \mathbb{R}$ sono differenziabili in un punto interno x_0 di A, allora cf, $f + g$, $f \cdot g$, e f/g se $g(x_0) \neq 0$, sono differenziabili in x_0 e*

- $\mathbf{D}(cf)(x_0) = c\mathbf{D}f(x_0)$,
- $\mathbf{D}(f + g)(x_0) = \mathbf{D}f(x_0) + \mathbf{D}g(x_0)$,
- $\mathbf{D}(f \cdot g)(x_0) = g(x_0)\mathbf{D}f(x_0) + f(x_0)\mathbf{D}g(x_0)$,
- $\mathbf{D}(f/g)(x_0) = \dfrac{g(x_0)\mathbf{D}f(x_0) - f(x_0)\mathbf{D}g(x_0)}{g(x_0)^2}$.

Per funzioni a valori vettoriali, guardando alle componenti o direttamente dalla definizione e utilizzando la Proposizione 10.9, è immediato verificare che

10.10 Proposizione. *Siano $f, g : A \subset \mathbb{R}^n \to \mathbb{R}^m$. Allora f, g sono differenziabili in x_0 se e solo se tutte le componenti f^i e g^i, $i = 1, m$ di f e g sono differenziabili in x, conseguentemente per ogni $\alpha, \beta \in \mathbb{R}$*

$$\frac{\partial(\alpha\, f + \beta\, g)^i}{\partial x^j}(x) = \alpha \frac{\partial f^i}{\partial x^j}(x) + \beta \frac{\partial g^i}{\partial x^j}(x)$$

o, in termini matriciali,

$$\mathbf{D}(\alpha\, f + \beta\, g)(x) = \alpha\, \mathbf{D}f(x) + \beta\, \mathbf{D}g(x).$$

Analogamente, se $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ e $\lambda : A \subset \mathbb{R}^n \to \mathbb{R}$ sono differenziabili, per ogni $j = 1, n$ e $i = 1, m$

$$\frac{\partial(\lambda f)^i}{\partial x^j}(x) = \lambda(x) \frac{\partial f^i}{\partial x^j}(x) + f^i(x) \frac{\partial \lambda}{\partial x^j}(x)$$

o, in termini matriciali,

$$\mathbf{D}(\lambda f)(x) = \lambda(x)\mathbf{D}f(x) + f(x)\mathbf{D}\lambda(x).$$

Si noti che $f(x) \in \mathbb{R}^m$ è un vettore colonna e $D\lambda(x_0)$ è un vettore riga a n elementi, e quindi $f(x)\mathbf{D}\lambda(x)$ è una matrice $m \times n$.

10.e Differenziale di funzioni composte

Il teorema del differenziale composto afferma che l'applicazione lineare tangente della composizione di due mappe differenziabili è la composizione delle applicazioni lineari tangenti delle due mappe. Precisamente

10.11 Teorema. *Siano $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ aperti, e siano $f : U \to \mathbb{R}^m$ con $f(U) \subset V$ e $g : V \to \mathbb{R}^p$ due funzioni. Se f è differenziabile in $x_0 \in U$ ed g è differenziabile in $f(x_0)$, allora $g \circ f$ è differenziabile in x_0 e*

$$\mathbf{D}(g \circ f)(x_0) = \mathbf{D}g(f(x_0))\mathbf{D}f(x_0) \tag{10.6}$$

righe per colonne.

Dimostrazione. Essendo $f(U) \subset V$, la funzione $g \circ f : U \to \mathbb{R}^p$ è ben definita. Ora per ipotesi, per $x \to x_0$ e $y \to f(x_0)$

$$\begin{aligned} f(x) &= f(x_0) + \mathbf{D}f(x_0)(x - x_0) + o(|x - x_0|) \qquad \text{per } x \to x_0 \\ g(y) &= g(f(x_0)) + \mathbf{D}g(f(x_0))(y - f(x_0)) + o(|y - f(x_0)|), \qquad \text{per } y \to f(x_0), \end{aligned}$$

da cui ricaviamo per $x \to x_0$

$$\begin{aligned} g(f(x)) &= g(f(x_0)) + \mathbf{D}g(f(x_0))\mathbf{D}f(x_0)(x - x_0) + \mathbf{D}g(f(x_0))o(|x - x_0|) \\ &\quad + o(|\mathbf{D}f(x_0)(x - x_0)|) + o(|x - x_0|) \\ &= g(f(x_0) + \mathbf{D}g(f(x_0))\mathbf{D}f(x_0)(x - x_0) + o(|x - x_0|). \end{aligned}$$

□

Il teorema è l'estenzione naturale al caso delle mappe vettoriali della (9.9) per il calcolo della derivata di una funzione lungo una curva. Si noti che la differenziabilità di g è in generale necessaria per la validità della (10.6), cfr. l'Esercizio 9.11.

10.12 Regola della catena. Supponiamo che $y : A \subset \mathbb{R}^m$ e $g : \mathbb{R}^m \to \mathbb{R}$ siano differenziabili rispettivamente in x_0 e $y(x_0)$. Se $y(x) = (y^1(x), \dots, y^m(x))^T$, esplicitando il prodotto righe per colonne in (10.6) si ha

$$\begin{aligned} \frac{\partial g \circ y}{\partial x^j}(x) &= \frac{\partial}{\partial x^j}\Big[g(y^1, y^2, \dots, y^m)\Big] = \mathbf{D}g(y(x))\mathbf{D}y(x) \\ &= \sum_{i=1}^m \frac{\partial g}{\partial y^i}\frac{\partial y^i}{\partial x^j} = \frac{\partial g}{\partial y^1}\frac{\partial y^1}{\partial x^j} + \frac{\partial g}{\partial y^2}\frac{\partial y^2}{\partial x^j} + \dots + \frac{\partial g}{\partial y^m}\frac{\partial y^m}{\partial x^j}. \end{aligned}$$

per $j = 1, \dots, m$; naturalmente è inteso che

$$\frac{\partial g}{\partial y^i} := \frac{\partial g}{\partial y^i}(y(x)), \qquad \text{e} \qquad \frac{\partial y^i}{\partial x^j} := \frac{\partial y^i}{\partial x^j}(x).$$

È la cosiddetta *regola della catena* per il calcolo delle derivate composte.

10.13 Esercizio. Siano $r : [0, 1] \to \mathbb{R}^m$ e $g : \mathbb{R}^m \to \mathbb{R}^p$, e sia $s(t) := g(r(t))$. Interpretare in questo caso la regola della catena.
Soluzione. Si ha

$$s'(t) = \mathbf{D}g(r(t))\, r'(t),$$

in particolare, supponendo che $r'(t_0)$ e $s'(t_0)$ non siano nulli, la matrice jacobiana di g manda la retta tangente in t_0 alla curva $r(t)$ sulla retta tangente alla curva immagine $s(t)$ in t_0.

10.14 Esercizio. La mappa $\phi :]0, +\infty[\times[0, 2\pi[\times[0, \pi] \to \mathbb{R}^3$ $\phi(\rho, \theta, \varphi) = (\rho\cos\theta\sin\varphi, \rho\sin\theta\sin\varphi, \rho\cos\varphi)$ è la mappa delle *coordinate polari* o *sferiche* in $\mathbb{R}^3$. Provare che

$$\mathbf{D}\phi(\rho, \theta, \varphi) = \begin{pmatrix} \cos\theta\sin\varphi & -\rho\sin\theta\sin\varphi & \rho\cos\theta\cos\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & \rho\sin\theta\cos\varphi \\ \cos\varphi & 0 & -\rho\sin\varphi \end{pmatrix}.$$

10.f Funzioni a valori matrici

Il calcolo per funzioni $t \to \mathbf{A}(t)$ a valori matrici $m \times n$ può essere ricondotto al calcolo per curve in $\mathbb{R}^{mn}$. Tuttavia per la rilevanza che esso ha, conviene sviluppare un certo numero di formule. Procedendo sulle componenti si provano le formule

$$\begin{aligned} (\mathbf{A}(t)x(t))' &= \mathbf{A}'(t)x(t) + \mathbf{A}(t)x'(t), \\ (\mathbf{A}(t)\mathbf{B}(t))' &= \mathbf{A}'(t)\mathbf{B}(t) + \mathbf{A}(t)\mathbf{B}'(t), \qquad (10.7) \\ (\lambda(t)\mathbf{A}(t))' &= \lambda'(t)\mathbf{A}(t) + \lambda(t)\mathbf{A}'(t), \\ (\operatorname{tr}\mathbf{A}(t))' &= \operatorname{tr}\mathbf{A}'(t), \qquad (10.8) \\ \frac{d}{dt}\, x(t) \bullet y(t) \;&=\; x'(t) \bullet y(t) \;+\; x(t) \bullet y'(t)\,. \end{aligned}$$

Inoltre

(i) partendo da $\mathbf{A}(t)\mathbf{A}(t)^{-1} = \mathrm{Id}$, si ricava dalla (10.7) che, se $\mathbf{A}(t)$ è differenziabile e invertibile, allora $\mathbf{A}(t)^{-1}$ è derivabile e
$$(\mathbf{A}(t)^{-1})' = -\mathbf{A}(t)^{-1}\mathbf{A}'(t)\mathbf{A}(t)^{-1}. \qquad (10.9)$$
(ii) Per induzione, dalle (10.7) e (10.8) $D(\operatorname{tr}\mathbf{A}(t)^2) = \operatorname{tr}(2\mathbf{A}(t)\mathbf{A}'(t))$, $D(\operatorname{tr}\mathbf{A}(t)^3) = \operatorname{tr}(3\mathbf{A}(t)^2\mathbf{A}'(t))$, …, $D(\operatorname{tr}\mathbf{A}(t)^n) = \operatorname{tr}(n\mathbf{A}(t)^{n-1}\mathbf{A}'(t))$. Perciò se p è un polinomio, si ha sempre
$$\frac{d}{dt}\Big(\operatorname{tr} p(\mathbf{A}(t)\Big) = \operatorname{tr}\Big(p'(\mathbf{A}(t))\,\mathbf{A}'(t)\Big). \qquad (10.10)$$
(iii) Se $\mathbf{A}(t)$ e $\mathbf{A}'(t)$ commutano fra loro, allora $D(\mathbf{A}(t)^n) = n\mathbf{A}(t)^{n-1}\mathbf{A}'(t)$ $\forall n$ e quindi per ogni polinomio p si ha
$$D(p(\mathbf{A}(t))) = p'(\mathbf{A}(t))\,\mathbf{A}'(t). \qquad (10.11)$$
Si noti che la regola della derivazione composta non vale in generale, ad esempio $(\mathbf{A}(t)^2)' = \mathbf{A}(t)\mathbf{A}'(t) + \mathbf{A}'(t)\mathbf{A}(t) \neq 2\mathbf{A}(t)\mathbf{A}'(t)$ se $\mathbf{A}(t)$ e $\mathbf{A}'(t)$ non commutano tra loro.
(iv) Osservando che il determinante è una funzione multilineare nelle sue colonne si ricava che per $\mathbf{A}(t) = [A_1(t) \,|\, A_2(t) \,|\, \dots \,|\, A_n(t)] \in M_{n,n}$ si ha
$$\begin{aligned}\frac{d}{dt}\det\mathbf{A}(t) = \det[A_1' \,|\, A_2 \,|\, \dots \,|\, A_n] \\ + \det[A_1 \,|\, A_2' \,|\, \dots \,|\, A_n] + \dots + \det[A_1 \,|\, A_2 \,|\, \dots \,|\, A_n'].\end{aligned}$$
Quindi, se $\mathbf{A}(0) = \mathrm{Id}$, si ha

$$\det[A_1'(0) \,|\, A_2(0) \,|\, \dots \,|\, A_n(0)] = (A_1^1)'(0),$$
$$\dots,$$
$$\det[A_1(0) \,|\, A_2(0) \,|\, \dots \,|\, A_n'(0)] = (A_n^n)'(0),$$

da cui

$$\frac{d \det \mathbf{A}(t)}{dt}(0) = \operatorname{tr} \mathbf{A}'(0) \qquad \text{se} \qquad \mathbf{A}(0) = \mathrm{Id}. \tag{10.12}$$

Se ora la matrice $\mathbf{Y}(s) \in M_{n,n}$ è invertibile per $s = t$, si ricava dalla (10.12) una formula per la variazione percentuale del determinante di un campo,

$$\frac{1}{\det \mathbf{Y}(t)} \frac{d \det \mathbf{Y}(s)}{ds}(t) = \frac{d \det(\mathbf{Y}(t)^{-1}\, \mathbf{Y}(s))}{ds}(t) = \operatorname{tr}\Big(\mathbf{Y}(t)^{-1}\, \mathbf{Y}'(t)\Big). \tag{10.13}$$