Università degli Studi di Firenze
Scuola di Scienze Matematiche, Fisiche e Naturali
C.d.L. Magistrale in Matematica

Anno Accademico 2013-2014
Tesi di Laurea

COMPLESSITÀ DELL'ALGORITMO DI STRASSEN PER IL PRODOTTO DI MATRICI. UN APPROCCIO TENSORIALE

Complexity of Strassen's algorithm for matrix multiplication. A tensor-based approach

Candidato:
Luca Simi

Relatore:
Prof. Giorgio Ottaviani

Indice

Introduzione

Nel 1969 il matematico tedesco Volker Strassen propose in [14] un algoritmo innovativo per calcolare il prodotto di matrici. La novità consisteva nelle prestazioni: era in grado di moltiplicare matrici $n \times n$ usando soltanto $O\left(n^{\log_2(7)}\right) \approx O\left(n^{2.81}\right)$ operazioni, contro le $O\left(n^3\right)$ dell'algoritmo "righe per colonne". Parte centrale dell'algoritmo è la possibilità di moltiplicare matrici 2×2 usando soltanto sette prodotti invece che otto, come verrà esposto in quanto segue.

Siano A e B matrici 2×2 e sia $C = AB$

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{bmatrix}$$

1. Calcoliamo i seguenti 7 prodotti:

$$\begin{aligned}
M_1 &= (A_{1,1} + A_{2,2})(B_{1,1} + B_{2,2}) \\
M_2 &= A_{1,1}(B_{1,2} - B_{2,2}) \\
M_3 &= A_{2,2}(B_{2,1} - B_{1,1}) \\
M_4 &= (A_{2,1} + A_{2,2})B_{1,1} \\
M_5 &= (A_{1,2} + A_{1,1})B_{2,2} \\
M_6 &= (A_{2,1} - A_{1,1})(B_{1,1} + B_{1,2}) \\
M_7 &= (A_{1,2} - A_{2,2})(B_{2,2} + B_{2,1})
\end{aligned}$$

2. Calcoliamo i coefficienti di C usando le seguenti combinazioni lineari

degli M_i:

$$\begin{aligned} C_{1,1} &= M_1 + M_3 - M_5 + M_7 \\ C_{1,2} &= M_2 + M_5 \\ C_{2,1} &= M_4 + M_3 \\ C_{2,2} &= M_1 - M_4 + M_2 + M_6 \end{aligned}$$

La validità dell'algoritmo è facilmente verificabile mediante il calcolo esplicito. Dal momento che il prodotto di matrici si comporta sui blocchi come sui coefficienti possiamo estendere l'algoritmo a matrici $2^k \times 2^k$ utilizzando il metodo "divide et impera": ad ogni iterazione si considerano A e B come matrici a blocchi 2×2 e si calcolano i sette prodotti in maniera ricorsiva. Siamo in grado di calcolare C usando in totale 7^k prodotti. Nel caso di matrici $n \times n$ basta aumentare il formato fino alla potenza di due più vicina, inserendo zero nei nuovi coefficienti e procedendo come nel caso già visto.

Protagonista nelle applicazioni computazionali, il prodotto di matrici cattura l'interesse del mondo tecnico e scientifico. La possibilità di migliorare gli algoritmi di cui disponiamo, oltre a costituire un argomento interessante dal punto di vista teorico, offre vantaggi pratici ed economici non indifferenti. In queste pagine saranno presentati argomenti volti a limitare dal basso il numero di prodotti che un algoritmo ottimale può impiegare. Vedremo come l'algoritmo di Strassen costituisca per il caso 2×2 un algoritmo ottimale (risultato provato indipendentemente in [15] e [8]), risultato per il quale verrà presentata una dimostrazione alternativa. Molti argomenti verranno accompagnati da alcuni script utilizzabili in maniera interattiva durante una sessione di *Macaulay2*.[1]

Vorrei rivolgere un ringraziamento particolare al Prof. Giorgio Ottaviani per l'argomento interessante, per i suggerimenti e le esperienze formative proposte durante il periodo dedicato alla tesi.

[1]Macaulay2 è un software dedicato alla ricerca in geometria algebrica e in algebra commutativa. Viene rilasciato sotto la licenza GNU GPL ed è liberamente scaricabile da `http://www.math.uiuc.edu/Macaulay2/`

Capitolo 1

Algebra Multilineare

L'algebra multilineare è lo studio delle applicazioni multilineari, ovvero lineari in ogni loro argomento. Il primo strumento che verrà presentato in questo capitolo è il linguaggio dei tensori, che permette di parlare in maniera generale e unificata degli oggetti dell'algebra lineare. Contemporaneamente fornirà un metodo preciso e geometrico per parlare di algoritmi. Le definizioni in questo capitolo costituiscono un adattamento di quanto contenuto in [2], in cui la trattazione è svolta in un contesto più generale. In particolare lo studio sarà rivolto a spazi vettoriali di dimensione finita.

1.1 Prodotti Tensoriali

Definizione. Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Il *prodotto tensoriale* di $V_1, \ldots, V_n$ è l'insieme

$$V_1 \otimes \ldots \otimes V_n = \{T : V_1^* \times \ldots \times V_n^* \longrightarrow \mathbb{K} \text{ multilineare}\}$$

Gli elementi di $V_1 \otimes \ldots \otimes V_n$ si dicono *tensori.* Inoltre se $d_i = \dim(V_i)$ diremo che $T \in V_1 \otimes \ldots \otimes V_n$ è un tensore di *ordine* n e *tipo* $d_1 \times \ldots \times d_n$.

Osservazione. In quanto segue saranno trattati esclusivamente prodotti tensoriali di spazi vettoriali di dimensione finita. In questo modo sarà lecito identificare uno spazio vettoriale V col suo biduale V^{**}. Tuttavia l'identificazione non è possibile per spazi vettoriali di dimensione infinita: se V è uno spazio vettoriale di dimensione infinita la cardinalità di V^* è sempre maggiore della cardinalità di V. Pertanto possiamo pensare ai tensori

$T \in V_1^* \otimes \ldots \otimes V_n^*$ come alle applicazioni multilineari

$$T : V_1 \times \ldots \times V_n \longrightarrow \mathbb{K}$$

Definiamo sul prodotto tensoriale le seguenti operazioni: per ogni $S, T \in V_1 \otimes \ldots \otimes V_n$ e $\lambda \in \mathbb{K}$ siano

$$\begin{aligned} (T+S)(\alpha_1, \ldots, \alpha_n) &= T(\alpha_1, \ldots, \alpha_n) + S(\alpha_1, \ldots, \alpha_n) \\ (\lambda T)(\alpha_1, \ldots, \alpha_n) &= \lambda T(\alpha_1, \ldots, \alpha_n) \end{aligned}$$

Con queste operazioni il prodotto tensoriale è uno spazio vettoriale.

Tra tutti i tensori ce ne sono alcuni più semplici da costruire e da calcolare, i tensori "decomponibili", definiti da n valutazioni sui V_i. Vedremo nella Proposizione 1.1.2 che ogni tensore in $V_1 \otimes \ldots \otimes V_n$ può essere scritto come somma di tensori decomponibili.

Definizione. Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Un tensore $T \in V_1 \otimes \ldots \otimes V_n$ si dice *decomponibile* se esistono $v_1 \in V_1, \ldots, v_n \in V_n$ tali che

$$T(\alpha_1, \ldots, \alpha_n) = \alpha_1(v_1) \cdot \ldots \cdot \alpha_n(v_n)$$

per ogni $\alpha_1 \in V_1^*, \ldots, \alpha_n \in V_n^*$. In questo caso indicheremo T con la notazione $v_1 \otimes \ldots \otimes v_n$.

Osservazione. Un tensore decomponibile in $V_1 \otimes \ldots \otimes V_n$ può essere definito usando diverse n-uple di vettori $v_1, \ldots, v_n$. La seguente proposizione specifica esattamente tale ambiguità.

Proposizione 1.1.1. *Sia $\mathbb{K}$ un campo, e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Siano $v_1, u_1 \in V_1, \ldots, v_n, u_n \in V_n$ vettori non nulli, allora $v_1 \otimes \ldots \otimes v_n = u_1 \otimes \ldots \otimes u_n$ se e solo se esistono scalari $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tali che $u_i = \lambda_i v_i$ per ogni i e $\lambda_1 \cdot \ldots \cdot \lambda_n = 1$.*

Dimostrazione. Procediamo per induzione su n.

- $n = 1$: $u_1 = v_1 \implies \lambda_1 = 1$
- $n > 1$: Supponiamo la tesi vera per $n-1$ coppie di vettori assegnati. Fissiamo $\beta \in V_n^*$ tale che $\beta(v_n) \neq 0$ e $\beta(u_n) \neq 0$, abbiamo

$$v_1 \otimes \ldots \otimes (v_{n-1}\beta(v_n)) = u_1 \otimes \ldots \otimes (u_{n-1}\beta(u_n))$$

Per ipotesi induttiva esistono $\mu_1, \ldots, \mu_{n-1}$ tali che $u_i = \mu_i v_i$ per $1 \leq i \leq n-2$, $\beta(u_n)u_{n-1} = \mu_{n-1}\beta(v_n)v_{n-1}$ e $\mu_1 \cdot \ldots \cdot \mu_{n-1} = 1$. Sostituendo otteniamo

$$v_1 \otimes \ldots \otimes v_{n-1} \otimes v_n = (\mu_1 v_1) \otimes \ldots \otimes \left(\mu_{n-1} \frac{\beta(v_n)}{\beta(u_n)} v_{n-1} \right) \otimes u_n$$

Dunque possiamo scrivere

$$v_1 \otimes \ldots \otimes v_{n-1} \otimes \left(v_n - \frac{\beta(v_n)}{\beta(u_n)} u_n \right) = 0$$

Pertanto segue

$$v_n = \frac{\beta(v_n)}{\beta(u_n)} u_n$$

Ponendo $\lambda_i = \mu_i$ per $1 \leq i \leq n-2, \lambda_{n-1} = \mu_{n-1}\frac{\beta(v_n)}{\beta(u_n)}, \lambda_n = \frac{\beta(u_n)}{\beta(v_n)}$ la tesi è provata.

□

Proposizione 1.1.2. *Sia $\mathbb{K}$ un campo, e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $\{v^k_1, \ldots, v^k_{d_k}\}$ base di V_k per $1 \leq k \leq n$, allora $V_1 \otimes \ldots \otimes V_n$ è uno spazio vettoriale con base $\{v^1_{i_1} \otimes \ldots \otimes v^n_{i_n}\}_{1 \leq i_k \leq d_k}$. Pertanto $dim(V_1 \otimes \ldots \otimes V_n) = \prod_{k=1}^n dim(V_k)$*

Dimostrazione. Sia $\{\alpha^k_1, \ldots, \alpha^k_{d_k}\}$ la base duale di $\{v^k_1, \ldots, v^k_{d_k}\}$ per $1 \leq k \leq n$, sia $T \in V_1 \otimes \ldots \otimes V_n$. Per ogni scelta di elementi $\beta^i \in V_i^*$ possiamo scrivere $\beta^i = \sum_{j_1} \lambda^i_{j_i} \alpha^i_{j_i}$ per qualche coefficiente $\lambda^i_{j_i} = \beta^i(v^i_{j_i})$, dunque:

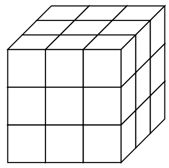
$$\begin{aligned} T(\beta^1, \ldots, \beta^n) &= T\left(\sum_{j_1} \lambda^1_{j_1} \alpha^1_{j_1}, \ldots, \sum_{j_n} \lambda^n_{j_n} \alpha^n_{j_n} \right) \\ &= \sum_{j_1, \ldots, j_n} \lambda^1_{j_1} \cdot \ldots \cdot \lambda^n_{j_n} T(\alpha^1_{j_1}, \ldots, \alpha^n_{j_n}) \\ &= \sum_{j_1, \ldots, j_n} \beta^1(v^1_{j_1}) \cdot \ldots \cdot \beta^n(v^n_{j_n}) T(\alpha^1_{j_1}, \ldots, \alpha^n_{j_n}) \\ &= \sum_{j_1, \ldots, j_n} T(\alpha^1_{j_1}, \ldots, \alpha^n_{j_n})(v^1_{j_1} \otimes \ldots \otimes v^n_{j_n})(\beta^1, \ldots, \beta^n) \end{aligned}$$

Dunque, ponendo $T_{j_1, \ldots, j_n} = T(\alpha^1_{j_1}, \ldots, \alpha^n_{j_n}) \in \mathbb{K}$, possiamo scrivere:

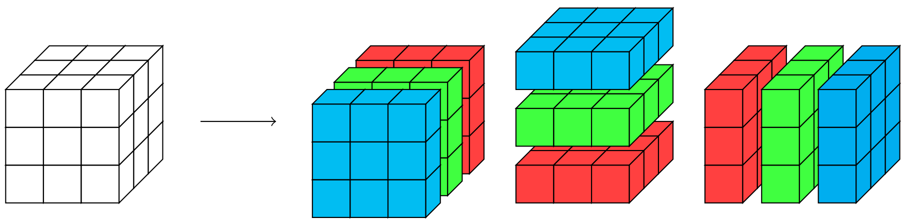
$$T = \sum_{j_1, \ldots, j_n} T_{j_1, \ldots, j_n} (v^1_{j_1} \otimes \ldots \otimes v^n_{j_n})$$

Osserviamo infine che $T_{i_1,\ldots,i_n} = T(\alpha^1_{i_1}, \ldots, \alpha^n_{i_n})$, dunque se $T = 0$ abbiamo $T_{i_1,\ldots,i_n} = 0$ per ogni $i_1, \ldots, i_n$. Dunque la tesi è provata. □

Osservazione. I tensori decomponibili costituiscono dunque gli ingredienti fondamentali per costruire un qualsiasi tensore, pertanto possiamo rappresentare un tensore come un array n-dimensionale (confrontare con il teorema 1.1.2):



Un modo equivalente per descrivere un tensore T nelle basi assegnate è usare una lista di *slice*: tensori di ordine inferiore ottenuti dalle sezioni di T lungo una direzione fissata.



1.2 Cambiamenti di Base

Definizione. Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Il *gruppo dei cambiamenti di base* è il gruppo

$$G = \mathrm{GL}(V_1) \times \ldots \times \mathrm{GL}(V_n)$$

che agisce sui tensori decomponibili in modo naturale ponendo

$$(g_1, \ldots, g_n)(v_1 \otimes \ldots \otimes v_n) = (g_1 v_1) \otimes \ldots \otimes (g_n v_n)$$

ed estendendo l'azione linearmente.

Osservazione. Fissate delle basi sugli spazi $V_1, \ldots, V_n$ un tensore è identificato da un array n-dimensionale. È interessante capire come questa rappresentazione si trasformi sotto cambiamenti di base negli spazi $V_1, \ldots, V_n$.

Proposizione 1.2.1. *Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $\{v_1^k, \ldots, v_{d_k}^k\}$ base di V_k, con base duale $\{\alpha_1^k, \ldots, \alpha_{d_k}^k\}$. Sia $T \in V_1 \otimes \ldots \otimes V_n$ e per $1 \leq k \leq n$ con coordinate $T_{i_1,\ldots,i_n}$ nella base corrispondente su $V_1 \otimes \ldots \otimes V_n$. Se $\{w_1^k, \ldots, w_{d_k}^k\}$ è una nuova base di V_k e vale $v_j^k = \sum_i a_{i,j}^k w_i^k$ le coordinate di T nella nuova base su $V_1 \otimes \ldots \otimes V_n$ sono date dagli scalari*

$$T'_{j_1,\ldots,j_n} = \sum_{i_1,\ldots,i_n} a^1_{j_1,i_1} \cdot \ldots \cdot a^n_{j_n,i_n} T_{i_1,\ldots,i_n}$$

Dimostrazione. Come già osservato nella Proposizione 1.1.2 possiamo scrivere

$$T_{j_1,\ldots,j_n} = T(\alpha^1_{j_1}, \ldots, \alpha^n_{j_n})$$

dunque

$$\begin{aligned}
T &= \sum_{i_1,\ldots,i_n} T_{i_1,\ldots,i_n} (v_{i_1} \otimes \ldots \otimes v_{i_n}) \\
&= \sum_{i_1,\ldots,i_n} T_{i_1,\ldots,i_n} \left[\left(\sum_{j_1} a^1_{j_1,i_1} w^1_{j_1} \right) \otimes \ldots \otimes \left(\sum_{j_n} a_{j_n,i_n} w^n_{j_n} \right) \right] \\
&= \sum_{j_1,\ldots,j_n} \left[\sum_{i_1,\ldots,i_n} a^1_{j_1,i_1} \cdot \ldots \cdot a^n_{j_n,i_n} T_{i_1,\ldots,i_n} \right] (w^1_{j_1} \otimes \ldots \otimes w^n_{j_n})
\end{aligned}$$

□

1.3 Proprietà del Prodotto Tensoriale

Proposizione 1.3.1. *Sia $\mathbb{K}$ un campo e siano U, V, W spazi vettoriali di dimensione finita su $\mathbb{K}$, allora:*

1. $U \otimes V \simeq V \otimes U$
2. $U \otimes V \otimes W \simeq (U \otimes V) \otimes W \simeq U \otimes (V \otimes W)$
3. *Se* $U \simeq U'$ *e* $V \simeq V'$ *allora* $U \otimes V \simeq U' \otimes V'$
4. $(U \otimes V)^* \simeq U^* \otimes V^*$.
5. $U^* \otimes V \simeq Hom(U, V)$

Dimostrazione. Per ogni punto dell'enunciato costruiamo esplicitamente un isomorfismo naturale (indipendente dalle basi assegnate).

1. Costruiamo l'isomorfismo mandando un tensore $A \in U \otimes V$ nel tensore $B \in V \otimes U$ definito ponendo per ogni $\alpha \in U^*, \beta \in V^*$

   $$B(\beta, \alpha) = A(\alpha, \beta)$$

   L'applicazione così definita è lineare e invertibile (basta invertire i ruoli di U e V nell'enunciato).

2. Definiamo l'isomorfismo sui tensori decomponibili mandando per ogni $u \in U, v \in V, w \in W$

   $$u \otimes v \otimes w \longmapsto (u \otimes v) \otimes w$$

   .

3. Siano $\phi : U \longrightarrow U'$ e $\psi : V \longrightarrow V'$ isomorfismi. Costruiamo l'isomorfismo mandando il tensore $A \in U \otimes V$ nel tensore $B \in U' \otimes V'$ definito da

   $$B(\alpha, \beta) = A(\alpha \circ \phi, \beta \circ \psi)$$

   per ogni $\alpha \in (U')^*, \beta \in (V')^*$. In questo modo è definita un'applicazione lineare invertibile (per costruire l'inversa è sufficiente usare ϕ^{-1} al posto di ϕ e ψ^{-1} al posto di ψ).

4. Costruiamo l'isomorfismo e la sua inversa. Per ogni $\alpha \in U^*, \beta \in V^*$ mandiamo $\alpha \otimes \beta$ in $L \in (U \otimes V)^*$ definito sui tensori decomponibili $u \otimes v$ da

   $$L(u \otimes v) = \alpha(u)\beta(v)$$

   ed estendendo linearmente.

   Viceversa per costruire l'inversa mandiamo $L \in (U \otimes V)^*$ nel tensore $G \in U^* \otimes V^*$ definito da

   $$G(u, v) = L(u \otimes v)$$

   per ogni $u \in U, v \in V$, ed estendendo bilinearmente.

5. Costruiamo l'isomorfismo e la sua inversa. Per ogni $A \in U^* \otimes V$ mandiamo A in $f : U \longrightarrow V$ definendo $f(u)$ per $u \in U$ come l'elemento di $V^{**} \simeq V$

   $$f(u)(\beta) = A(u, \beta)$$

per ogni $\beta \in V^*$.

Viceversa mandiamo $f \in \mathrm{Hom}(U,V)$ in $A \in U^* \otimes V$ definita ponendo

$$A(u,\beta) = \beta(f(u))$$

per ogni $u \in U, \beta \in V^*$.

□

1.4 Tensori Simmetrici e Tensori Alternanti

In questa sezione lavoreremo sempre in un campo $\mathbb{K}$ con caratteristica nulla. Consideriamo un tensore $T \in V^{\otimes n}$ e una permutazione $\sigma \in \mathfrak{S}_n$. Costruiamo $\sigma \circ T \in V^{\otimes n}$ ponendo:

$$\sigma \circ T : (\alpha_1, \dots, \alpha_n) \longmapsto T\left(\alpha_{\sigma(1)}, \dots, \alpha_{\sigma(n)}\right)$$

Definizione. Siano $\mathbb{K}$ un campo e V uno spazio vettoriale di dimensione finita su $\mathbb{K}$. Un tensore $T \in V^{\otimes n}$ si dice *simmetrico* se

$$\sigma \circ T = T$$

per ogni $\sigma \in \mathfrak{S}_n$. Indichiamo con S^nV l'insieme dei tensori simmetrici in $V^{\otimes n}$.

Definizione. Siano $\mathbb{K}$ un campo e V uno spazio vettoriale di dimensione finita su $\mathbb{K}$. Un tensore $T \in V^{\otimes n}$ si dice *alternante* se

$$\sigma \circ T = \mathrm{sgn}(\sigma) T$$

per ogni $\sigma \in \mathfrak{S}_n$. Indichiamo con Λ^nV l'insieme dei tensori alternanti in $V^{\otimes n}$

Osservazione. È facile verificare che S^nV e Λ^nV sono sottospazi vettoriali di $V^{\otimes n}$.

Definizione. Sia V uno spazio vettoriale di dimensione finita su $\mathbb{K}$. Definiamo *operatore di simmetrizzazione* l'applicazione lineare $\pi_S : V^{\otimes n} \longrightarrow V^{\otimes n}$ definita ponendo

$$\pi_S(T) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sigma \circ T$$

per ogni $T \in V^{\otimes n}$.

Definizione. Sia V uno spazio vettoriale di dimensione finita su $\mathbb{K}$. Definiamo *operatore di antisimmetrizzazione* l'applicazione lineare $\pi_\Lambda : V^{\otimes n} \longrightarrow V^{\otimes n}$ definita ponendo

$$\pi_\Lambda(T) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \cdot \sigma \circ T$$

per ogni $T \in V^{\otimes n}$.

Proposizione 1.4.1. *Sia V uno spazio vettoriale di dimensione finita su $\mathbb{K}$, allora*

- $\pi_S(T) \in S^n V$ *per ogni* $T \in V^{\otimes n}$
- $\pi_\Lambda(T) \in \Lambda^n V$ *per ogni* $T \in V^{\otimes n}$

Dimostrazione. Sia $\tau \in \mathfrak{S}_n$, allora

$$\begin{aligned}
\tau \circ \pi_S(T) &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} (\tau\sigma) \circ T \\
&= \frac{1}{n!} \sum_{(\tau\sigma) \in \mathfrak{S}_n} (\tau\sigma) \circ T \\
&= \frac{1}{n!} \sum_{\mu \in \mathfrak{S}_n} \mu \circ T
\end{aligned}$$

Analogamente

$$\begin{aligned}
\tau \circ \pi_\Lambda(T) &= \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma)(\tau\sigma) \circ T \\
&= \frac{1}{n!} \operatorname{sgn}(\tau) \sum_{(\tau\sigma) \in \mathfrak{S}_n} \operatorname{sgn}(\tau\sigma)(\tau\sigma) \circ T \\
&= \operatorname{sgn}(\tau) \frac{1}{n!} \sum_{\mu \in \mathfrak{S}_n} \operatorname{sgn}(\mu)\mu \circ T \\
&= \operatorname{sgn}(\tau)\pi_\Lambda(T)
\end{aligned}$$

□

Proposizione 1.4.2. *Sia V uno spazio vettoriale su $\mathbb{K}$, allora:*

- $\pi_S(T) = T$ *per ogni* $T \in S^n V$

- $\pi_S(T) = 0$ *per ogni* $T \in \Lambda^n V$
- $\pi_\Lambda(T) = 0$ *per ogni* $T \in S^n V$
- $\pi_\Lambda(T) = T$ *per ogni* $T \in \Lambda^n V$

Dimostrazione. Seguono immediatamente dalla definizione di tensori simmetrici e alternanti. □

Corollario 1.4.1. *Sia* V *uno spazio vettoriale su* $\mathbb{K}$*, allora:*

- $S^n V = \pi_S\left(V^{\otimes n}\right)$
- $\Lambda^n V = \pi_\Lambda\left(V^{\otimes n}\right)$

Dimostrazione. Seguono direttamente dalle Proposizioni 1.4.1 e 1.4.2. □

Proposizione 1.4.3. *Sia* V *uno spazio vettoriale su* $\mathbb{K}$ *di dimensione* d*, allora:*

- $dim\left(\Lambda^n V\right) = \binom{d}{n}$
- $dim\left(S^n V\right) = \binom{d+n-1}{n}$

Dimostrazione. Una base di $S^n V$ è costituita dai tensori $v_{i_1} \otimes \ldots \otimes v_{i_n}$ con $i_1 \leq \ldots \leq i_n$, mentre una base di $\Lambda^n V$ è costituita dai tensori $v_{i_1} \otimes \ldots \otimes v_{i_n}$ con $i_1 < \ldots < i_n$. Le dimensioni sono pertanto quelle enunciate. □

Proposizione 1.4.4. *Sia* V *uno spazio vettoriale su* $\mathbb{K}$*, allora* $V \otimes V = S^2 V \oplus \Lambda^2 V$*.*

Dimostrazione. Consideriamo un tensore $u \otimes v \in V \otimes V$, osserviamo:

$$u \otimes v = \frac{1}{2}\left(u \otimes v - v \otimes u\right) + \frac{1}{2}\left(u \otimes v + v \otimes u\right)$$

La decomposizione si estende al caso generale per linearità. □

1.5 Decomposizione di $V^{\otimes 3}$ in GL(V)-sottomoduli

Il gruppo GL(V) agisce sul prodotto tensoriale $V^{\otimes 3}$ in modo naturale. Con questa azione diciamo che $V^{\otimes 3}$ è un GL(V)-modulo. Chiameremo GL(V)-sottomodulo un sottospazio GL(V)-invariante di $V^{\otimes 3}$. Abbiamo visto come $V^{\otimes 2}$ sia decomponibile in $S^2V \oplus \Lambda^2 V$. Aggiungendo un fattore gli spazi S^3V e $\Lambda^3 V$ non sono sufficienti per ricostruire $V^{\otimes 3}$.

Definizione. Definiamo le seguenti applicazioni $\rho : V^{\otimes 3} \longrightarrow V^{\otimes 3}$

$$
\begin{aligned}
&\rho_{\ytableaushort{123}} = \pi_S \\
&\rho_{\ytableaushort{1,2,3}} = \pi_\Lambda \\
&\rho_{\ytableaushort{1,2}} : u \otimes v \otimes w \longmapsto \frac{1}{2} u \otimes v \otimes w - \frac{1}{2} v \otimes u \otimes w \\
&\rho_{\ytableaushort{13}} : u \otimes v \otimes w \longmapsto \frac{1}{2} u \otimes v \otimes w + \frac{1}{2} w \otimes v \otimes u \\
&\rho_{\ytableaushort{1,3}} : u \otimes v \otimes w \longmapsto \frac{1}{2} u \otimes v \otimes w - \frac{1}{2} w \otimes v \otimes u \\
&\rho_{\ytableaushort{12}} : u \otimes v \otimes w \longmapsto \frac{1}{2} u \otimes v \otimes w + \frac{1}{2} v \otimes u \otimes w \\
&\rho_{\ytableaushort{13,2}} = \rho_{\ytableaushort{13}} \circ \rho_{\ytableaushort{1,2}} \\
&\rho_{\ytableaushort{12,3}} = \rho_{\ytableaushort{12}} \circ \rho_{\ytableaushort{1,3}}
\end{aligned}
$$

Definendo $S_\rho V = \rho\left(V^{\otimes 3}\right)$ per ogni ρ come sopra otteniamo la seguente decomposizione di $V^{\otimes 3}$ in GL(V)-sottomoduli:

Proposizione 1.5.1. $V^{\otimes 3} = S_{\ytableaushort{123}} V \oplus S_{\ytableaushort{12,3}} V \oplus S_{\ytableaushort{13,2}} V \oplus S_{\ytableaushort{1,2,3}} V$

Dimostrazione. Svolgendo i calcoli per i tensori decomponibili di tipo $T = u \otimes v \otimes w$ vale

$$
T = \rho_{\ytableaushort{123}}(T) + \frac{4}{3}\rho_{\ytableaushort{12,3}}(T) + \frac{4}{3}\rho_{\ytableaushort{13,2}}(T) + \rho_{\ytableaushort{1,2,3}}(T)
$$

dunque

$$
V^{\otimes 3} = S_{\ytableaushort{123}} V + S_{\ytableaushort{12,3}} V + S_{\ytableaushort{13,2}} V + S_{\ytableaushort{1,2,3}} V
$$

Per verificare che la somma è diretta osserviamo che

$$
\begin{aligned}
\rho^2_{\ytableaushort{123}} &= \rho_{\ytableaushort{123}} \\
\rho^2_{\ytableaushort{13,2}} &= 3\rho_{\ytableaushort{13,2}} \\
\rho^2_{\ytableaushort{12,3}} &= 3\rho_{\ytableaushort{12,3}} \\
\rho^2_{\ytableaushort{1,2,3}} &= \rho_{\ytableaushort{1,2,3}}
\end{aligned}
$$

Dato uno spazio vettoriale U e un endomorfismo $f : U \longrightarrow U$ tale che esiste $\lambda \in \mathbb{K} \setminus \{0\}$ per cui $f^2 = \lambda f$ possiamo decomporre $U = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$, infatti per un generico $u \in U$ vale

$$u = f\left(\frac{1}{\lambda}u\right) + \left(u - f\left(\frac{1}{\lambda}u\right)\right)$$

Dunque per ogni funzione ρ_α tra quelle appena elencate possiamo scrivere $V^{\otimes 3} = \mathrm{Ker}(\rho_\alpha) \oplus \mathrm{Im}(\rho_\alpha)$. Dal momento che ogni spazio $S_\alpha V = \mathrm{Im}(\rho_\alpha)$ per provare che la somma è diretta è sufficiente verificare le seguenti relazioni

$$
\begin{aligned}
S_{\ytableaushort{12,3}}V + S_{\ytableaushort{13,2}}V + S_{\ytableaushort{1,2,3}}V &\subseteq \mathrm{Ker}\left(\rho_{\ytableaushort{123}}\right) \\
S_{\ytableaushort{123}}V + S_{\ytableaushort{13,2}}V + S_{\ytableaushort{1,2,3}}V &\subseteq \mathrm{Ker}\left(\rho_{\ytableaushort{12,3}}\right) \\
S_{\ytableaushort{123}}V + S_{\ytableaushort{12,3}}V + S_{\ytableaushort{1,2,3}}V &\subseteq \mathrm{Ker}\left(\rho_{\ytableaushort{13,2}}\right) \\
S_{\ytableaushort{123}}V + S_{\ytableaushort{13,2}}V + S_{\ytableaushort{12,3}}V &\subseteq \mathrm{Ker}\left(\rho_{\ytableaushort{1,2,3}}\right)
\end{aligned}
$$

Ovvero provando che ogni mappa ρ_α composta con una distinta ρ_β otteniamo la mappa nulla. Questo è facilmente verificabile sui tensori decomponibili attraverso il calcolo esplicito. $\square$

1.6 Rango

Definizione. Siano $\mathbb{K}$ un campo e $V_1, \ldots, V_n$ spazi vettoriali su $\mathbb{K}$. Dato un tensore $T \in V_1 \otimes \ldots \otimes V_n$ il *rango* di T è il minimo intero $R(T)$ tale che T è esprimibile come somma di $R(T)$ tensori decomponibili.

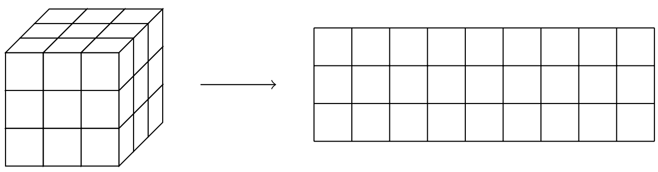
Come mostrato in [9] molti problemi associati allo studio dei tensori sono NP-hard, e viene usato il rango di un tensore come misura della sua complessità. In particolare calcolare il rango è un problema complesso: dato un tensore $T \in U \otimes V \otimes W$ e un intero r, determinare se $R(T) \leq r$ è un problema NP-completo (vedere [6] per una prova). Pertanto un approccio esaustivo diventa impraticabile anche per dimensioni non troppo elevate.

1.7 Flattening

Definizione. Siano $\mathbb{K}$ un campo e U, V, W spazi vettoriali su $\mathbb{K}$. Dato $T \in U \otimes V \otimes W$ il *flattening* di T su W è l'applicazione lineare T_W costruita considerando $T \in \mathrm{Hom}((U \otimes V)^*, W)$. Definiamo il flattening sui tensori decomponibili $u \otimes v \otimes w$ ponendo

$$(u \otimes v \otimes w)_W : \alpha \longmapsto \alpha(u \otimes v)w$$

ed esteso linearmente.



Proposizione 1.7.1. *Sia $T \in U \otimes V \otimes W$, allora $rk(T_W) \leq R(T)$.*

Dimostrazione. Possiamo scrivere T come somma di $R = R(T)$ tensori decomponibili

$$T = \sum_{r=1}^{R} u_r \otimes v_r \otimes w_r$$

dunque

$$T_W(\alpha \otimes \beta) = \sum_{r=1}^{R} \alpha(u_r)\beta(v_r)w_r$$

e $\dim(\mathrm{Im}(T_W)) \leq R$. □

Proposizione 1.7.2. *Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $T \in V_1 \otimes \ldots \otimes V_n$. Sia*

$$T_i : V_1^* \times \ldots \times V_{i-1}^* \times V_{i+1}^* \times \ldots \times V_n^* \longrightarrow V_i$$

definita da

$$T_i(\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n)(\alpha_i) = T(\alpha_1, \ldots, \alpha_n)$$

per $\alpha_i \in V_i^*$ *allora* $R(T) = 1$ *se e solo se* $rk(T_i) = 1$ *per ogni* $1 \leq i \leq n$.

Dimostrazione. Possiamo scrivere

$$T = \sum_{k=1}^{R(T)} v_1^{(k)} \otimes \ldots \otimes v_n^{(k)}$$

Dunque

$$T_i(\alpha_1, \ldots, \hat{\alpha_i}, \ldots, \alpha_n) = \sum_{k=1}^{R(T)} \alpha_1\left(v_1^{(k)}\right) \cdot \ldots \cdot \alpha_n\left(v_n^{(k)}\right) v_i^{(k)}$$

Pertanto $\mathrm{Im}(T_i) = \left\langle v_i^{(1)}, \ldots, v_i^{(R(T))} \right\rangle$. Possiamo affermare che se $R(T) = 1$ allora $\mathrm{rk}(T_i) = 1$ per ogni i. Viceversa se $\mathrm{rk}(T_i) = 1$ per ogni i allora esiste $v_i \in V_i$ tale che per qualche $\lambda_k \in \mathbb{K}$ vale $v_i^{(k)} = \lambda_i^{(k)} v_i$ per $1 \leq k \leq R(T)$. Dunque, sostituendo

$$\begin{aligned} T &= \sum_{k=1}^{R(T)} \left(\lambda_1^{(k)} v_1\right) \otimes \ldots \otimes \left(\lambda_n^{(k)} v_n\right) \\ &= \left[\left(\sum_{k=1}^{R(T)} \lambda_1^{(k)}\right) v_1\right] \otimes \ldots \otimes \left[\left(\sum_{k=1}^{R(T)} \lambda_n^{(k)}\right) v_n\right] \end{aligned}$$

Dunque $R(T) = 1$. □

Macaulay2. Riportiamo di seguito una funzione per calcolare il flattening di un tensore. Nel codice viene utilizzata la funzione *diff* [1], che in questo caso restituisce una matrice con i valori assunti dagli elementi del duale sul tensore.

```
-- input: t scritto come polinomio, e gli spazi U e V come ideali
-- output: t_U: U^* -> V scritto come matrice
Flatten = (t, U, V) -> (
     R := ring(t);
     K := coefficientRing(R);
     u := numgens(U);
     varU := gens(U);
     v := numgens(V);
     varV := gens(V);
```

[1] Per una descrizione più articolata consultare la pagina: `http://www.math.uiuc.edu/Macaulay2/doc/Macaulay2-1.6/share/doc/Macaulay2/Macaulay2Doc/html/_diff.html`

```
map(K^v, K^u, (i, j) -> (
          sub(diff(varV_(0, i), diff(varU_(0, j), t)), K)
          )
     )
)
```

Capitolo 2

Geometria Algebrica

La geometria algebrica è lo studio dei luoghi degli zeri di famiglie di polinomi. In questa sezione verranno introdotti gli elementi essenziali per ambientare lo studio dei tensori e del loro rango nel contesto della geometria algebrica. Gran parte delle definizioni sono tratte da [5].

2.1 Varietà Proiettive

Sia $\mathbb{K}$ un campo (algebricamente chiuso). Indichiamo con $\mathbb{P}^n$ lo spazio proiettivo su $\mathbb{K}^{n+1}$.

Definizione. Data una famiglia di polinomi omogenei $\mathcal{F} \subseteq \mathbb{K}[x_0, \dots, x_n]$ il *luogo degli zeri* di $\mathcal{F}$ è

$$Z(\mathcal{F}) = \{[P] \in \mathbb{P}^n \text{ t.c. } f(P) = 0 \text{ per ogni } f \in \mathcal{F}\}$$

Una *varietà proiettiva* è un insieme $X \subseteq \mathbb{P}^n$ per cui esiste una famiglia di polinomi omogenei $\mathcal{F} \subseteq \mathbb{K}[x_0, \dots, x_n]$ tale che $X = Z(\mathcal{F})$.

Osservazione. Se $f \in \mathbb{K}[x_0, \dots, x_n]$ è un polinomio omogeneo $f(\lambda P) = \lambda^{\deg(f)} f(P)$, dunque $f(P) = 0$ se e solo se $f(\lambda P) = 0$, pertanto la definizione di varietà proiettiva è ben posta.

Proposizione 2.1.1. *Siano $\mathcal{F}, \mathcal{G}, \mathcal{F}_\lambda \subseteq \mathbb{K}[x_0, \dots, x_n]$ famiglie di polinomi omogenei. Allora:*

- $Z(\mathcal{F}) \cup Z(\mathcal{G}) = Z\left(\{fg \textit{ t.c. } f \in \mathcal{F}, g \in \mathcal{G}\}\right)$

- $\bigcap_\lambda Z(\mathcal{F}_\lambda) = Z\left(\bigcup_\lambda \mathcal{F}_\lambda\right)$
- $\emptyset = Z(\{1\})$
- $\mathbb{P}^n = Z(\{0\})$

Osservazione. La proposizione appena vista mostra come le varietà proiettive soddisfano gli assiomi per insiemi chiusi, dunque ci consente di definire una topologia su $\mathbb{P}^n$.

Definizione. La *topologia di Zariski* su $\mathbb{P}^n$ è l'unica topologia definita scegliendo come chiusi le varietà proiettive.

Definizione. Una varietà $X \subseteq \mathbb{P}^n$ si dice *irriducibile* se per ogni coppia di varietà $Y, Z \subseteq X$ tale che $X = Y \cup Z$ si ha $Y = X$ oppure $Z = X$.

Definizione. Dato $X \subseteq \mathbb{P}^n$ possiamo definire l'*ideale omogeneo* di X come

$$I(X) = \{f \in \mathbb{K}[x_0, \dots, x_n] \text{ omogeneo t.c. } f(P) = 0 \text{ per ogni } [P] \in X\}$$

Osservazione. $I(X)$ è effettivamente un ideale omogeneo in $\mathbb{K}[x_0, \dots, x_n]$.

Definizione. Un anello R si dice *Noetheriano* se ogni suo ideale è finitamente generato.

Proposizione 2.1.2. *Sia R un anello. Sono equivalenti:*

1. *R è noetheriano.*
2. *Ogni catena ascendente infinita $I_1 \leq I_2 \leq \ldots \leq I_n \leq \ldots$ di ideali di R si stabilizza, cioè esiste $m \in \mathbb{N}$ tale che $I_m = I_k$ per ogni $k \geq m$.*
3. *Ogni famiglia $\emptyset \neq \mathcal{F}$ di ideali di R ha un elemento massimale.*

Vale il seguente risultato dovuto a Hilbert, che consente di studiare ideali nell'anello di polinomi come generati da una quantità finita di polinomi.

Teorema 2.1.1 (Basissatz)**.** *Sia R un anello. Se R è noetheriano allora $R[x]$ è noetheriano.*

Osservazione. Applicando il teorema induttivamente otteniamo che $\mathbb{K}[x_0, \dots, x_n]$ è un anello noetheriano. Pertanto possiamo affermare che $I(X)$ è finitamente generato, ovvero:

$$I(X) = (f_1, \dots, f_m)$$

per qualche polinomio omogeneo $f_1, \dots, f_m \in \mathbb{K}[x_0, \dots, x_n]$. Dunque possiamo affermare $Z(I(X)) = Z(\{f_1, \dots, f_n\})$, ovvero ogni varietà proiettiva può essere pensata come un sistema polinomiale finito. Oppure, geometricamente, come un intersezione finita di varietà definite da una sola equazione polinomiale (omogenea). Osserviamo inoltre che un chiuso di Zariski è anche un chiuso euclideo in quanto intersezione (finita) di retroimmagini di 0 mediante un polinomio.

Definizione. Una topologia su uno spazio X si dice *Noetheriana* se ogni catena discendente di chiusi

$$Y_1 \supseteq Y_2 \supseteq \dots \supseteq Y_n \supseteq \dots$$

si stabilizza, cioè se esiste $m \in \mathbb{N}$ tale che $Y_m = Y_k$ per ogni $k \geq m$.

Proposizione 2.1.3. *La topologia di Zariski su $\mathbb{P}^n$ è noetheriana.*

Dimostrazione. Supponiamo per assurdo che esista una catena discendente infinita di chiusi tutti distinti $Y_1 \supsetneq Y_2 \supsetneq \dots \supsetneq Y_n \supsetneq \dots$. Corrispondentemente abbiamo una catena ascendente di ideali tutti distinti

$$I(Y_1) \subsetneq I(Y_2) \subsetneq \dots \subsetneq I(Y_n) \subsetneq \dots$$

Dunque un assurdo, dal momento che $\mathbb{K}[x_0, \dots, x_n]$ è un anello noetheriano. □

Proposizione 2.1.4. *Sia $S \subseteq \mathbb{P}^n$. Indicando con $\overline{S}$ la chiusura di S nella topologia di Zariski abbiamo $\overline{S} = Z(I(S))$.*

Definizione. Sia $J \leq \mathbb{K}[x_0, \dots, x_n]$ un ideale. Il *radicale* di J è l'insieme

$$\sqrt{J} = \{f \in \mathbb{K}[x_0, \dots, x_n] \ t.c. \ f^k \in J \ per \ qualche \ k \in \mathbb{N}\}$$

L'ideale J di dice *radicale* se $J = \sqrt{J}$.

Proposizione 2.1.5. *Sia $J \leq \mathbb{K}[x_0, \dots, x_n]$ un ideale, allora $\sqrt{J}$ è un ideale. Inoltre se J è omogeneo anche $\sqrt{J}$ è omogeneo.*

Proposizione 2.1.6. *Sia $J \leq \mathbb{K}[x_0, \dots, x_n]$ un ideale, allora $Z(J) = Z(\sqrt{J})$.*

Il rapporto tra $Z(\circ)$ e $I(\circ)$ è descritto dal Teorema degli Zeri di Hilbert, che stabilisce una corrispondenza biunivoca tra varietà e ideali radicali.

Teorema 2.1.2 (Nullstellensatz). *Sia $\mathbb{K}$ un campo algebricamente chiuso e $J \leq \mathbb{K}[x_1, \dots, x_n]$ un ideale. Allora*

$$I(Z(J)) = \sqrt{J}$$

2.2 Varietà di Segre

Vediamo adesso un modo naturale per ambientare lo studio dei tensori decomponibili in uno spazio proiettivo.

Definizione. Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Definiamo *embedding di Segre* l'applicazione

$$\mathrm{Seg} : \mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n \longrightarrow \mathbb{P}(V_1 \otimes \ldots \otimes V_n)$$

ponendo

$$\mathrm{Seg}([v_1], \ldots, [v_n]) = [v_1 \otimes \ldots \otimes v_n]$$

Osservazione. Sia $d_i = dim(V_i)$. Assegnando delle basi negli spazi $V_1, \ldots, V_n$ l'embedding di Segre ha la forma

$$\mathrm{Seg} : ([v_1], \ldots, [v_n]) = [v_1^0 \cdot \ldots \cdot v_n^0, \ldots, v_1^{d_1} \cdot \ldots \cdot v_n^{d_n}]$$

dove $v_i^0, \ldots, v_i^{d_i}$ sono le componenti di v_i nella base assegnata.

Proposizione 2.2.1. *Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. L'embedding di Segre*

$$Seg : \mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n \longrightarrow \mathbb{P}(V_1 \otimes \ldots \otimes V_n)$$

è iniettivo.

Dimostrazione. L'enunciato segue immediatamente dalla Proposizione 1.1.1. □

Proposizione 2.2.2. *$Seg(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n)$ è una varietà proiettiva.*

Dimostrazione. Abbiamo visto nella Proposizione 1.7.2 che un tensore $T \in V_1 \otimes \ldots \otimes V_n$ ha rango $R(T) = 1$ se e solo se tutti i suoi flattening T_i hanno rango $\mathrm{rk}(T_i) = 1$. Dunque i polinomi (omogenei) che definiscono la varietà di Segre sono i minori 2×2 di tutti i flattening T_i. □

Osservazione. Il gruppo $G = \mathrm{GL}(V_1) \times \ldots \times \mathrm{GL}(V_k)$ dei cambiamenti di base in $V_1, \ldots, V_k$ agisce sulla varietà di Segre $X = \mathrm{Seg}(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_k)$ come segue:

$$(g_1, \ldots, g_k) : [v_1 \otimes \ldots \otimes v_k] \longmapsto [(g_1 v_1) \otimes \ldots \otimes (g_k v_k)]$$

La varietà di Segre è invariante sotto l'azione di G appena definita. Diremo che X è G-invariante.

2.3 Varietà Secanti

Abbiamo visto come sia possibile studiare i tensori decomponibili nel contesto della geometria algebrica. Adesso vediamo come estendere lo studio a tensori di rango qualsiasi.

Definizione. I punti $[P_1], \dots, [P_k] \in \mathbb{P}^n$ sono *in posizione generale* se $P_1, \dots, P_k$ sono linearmente indipendenti, oppure se $k > n+1$ e $n+1$ tra questi, comunque scelti, sono linearmente indipendenti.

Definizione. Siano $X_1, \dots, X_k \subseteq \mathbb{P}^n$ varietà proiettive. Il *join* di $X_1, \dots, X_n$ è la varietà

$$J(X_1, \dots, X_n) = \overline{\bigcup_{\substack{[P_1]\in X_1, \dots, [P_k]\in X_k \\ \text{in posizione generale}}} \langle [P_1], \dots, [P_k] \rangle}$$

Definizione. Sia $X \subseteq \mathbb{P}^n$ una varietà proiettiva. La *k-esima secante a X* è la varietà

$$\sigma_k(X) = J(\underbrace{X, \dots, X}_{\text{k volte}})$$

Osservazione. In questa definizione la chiusura è intesa nella topologia di Zariski, dunque possiamo pensare a $\sigma_k(X)$ come *la più piccola varietà proiettiva contenente sottospazi di dimensione (proiettiva) $k-1$ passanti per X.* Una prima osservazione è la seguente:

$$X = \sigma_1(X) \subseteq \sigma_2(X) \subseteq \dots \subseteq \sigma_k(X) \subseteq \dots \subseteq \mathbb{P}^n$$

2.4 Rango Bordo

Aver utilizzato una chiusura topologica nella definizione di varietà secanti ci porta a definire una grandezza leggermente diversa dal rango (e più topologica) per misurare la complessità di un tensore.

Definizione. Sia $\mathbb{K}$ un campo e siano $V_1, \dots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $X = \mathrm{Seg}(\mathbb{P}V_1 \times \dots \times \mathbb{P}V_n)$, sia $[T] \in \mathbb{P}(V_1 \otimes \dots \otimes V_n)$. Il *rango bordo* di T è il minimo intero $\underline{\mathrm{R}}(T)$ tale che $[T] \in \sigma_{\underline{\mathrm{R}}(T)}(X)$.

Osservazione. È possibile definire un tensore di rango bordo $\leq r$ come limite di una successione di tensori di rango $\leq r$.

Proposizione 2.4.1. *Sia $X \subseteq \mathbb{P}^n$ una varietà irriducibile e sia $\emptyset \neq U \subseteq X$ un aperto di Zariski, allora $\overline{U} = X$ sia come chiusura di Zariski sia come chiusura euclidea.*

Dimostrazione. Per la chiusura di Zariski: $X = (X \setminus U) \cup \overline{U}$. Poiché $U \neq \emptyset$ deve essere $X \setminus U \neq X$, dunque $\overline{U} = X$ per irriducibilità.

Per la chiusura euclidea: esistono polinomi omogenei $f_1, \ldots, f_k$ tali che $U = \{[P] \in X \text{ t.c. } f_1(P) \neq 0, \ldots, f_k(P) \neq 0\}$, dunque per la continuità dei polinomi $f_1, \ldots, f_k$ la chiusura di U contiene anche i punti di X in cui gli f_i si annullano, dunque $\overline{U} = X$. □

Vediamo nel seguente corollario come sia possibile definire il rango bordo in termini di limiti nella topologia euclidea. Per una trattazione più approfondita consultare [10] (Corollario 5.1.1.5).

Corollario 2.4.1. *Siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{C}$. Sia $X = Seg(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n)$, sia $[T] \in \mathbb{P}(V_1 \otimes \ldots \otimes V_n)$, allora $\underline{R}(T) \leq r$ se e solo se esiste una successione di tensori $\{T_k\}_{k \in \mathbb{N}} \subseteq V_1 \otimes \ldots \otimes V_n$ di rango $R(T_k) \leq r$ convergente a T nella topologia euclidea.*

Dimostrazione. (Cenni) $\sigma_r(X)$ è una varietà irriducibile (perché X è irriducibile e il join di due varietà irriducibili è irriducibile). Sia U l'insieme degli elementi $[T]$ per cui $R(T) \leq r$. Ponendo nella Proposizione 2.4.1 $Z = \sigma_r(X)$ il corollario è provato. □

Osservazione. Possiamo parlare di *algoritmo approssimato* per un tensore T quando consideriamo una successione T_n di tensori convergente a T e consideriamo per ogni n una decomposizione di T_n di lunghezza $\underline{R}(T)$.

Proposizione 2.4.2. *Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $[T] \in \mathbb{P}(V_1 \otimes \ldots \otimes V_n)$ allora $\underline{R}(T) \leq R(T)$.*

Dimostrazione. Se $X = \text{Seg}(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n)$ allora $T \in \sigma_{R(T)}(X)$ per definizione. □

Proposizione 2.4.3. *Sia $\mathbb{K}$ un campo, sia $M \in \mathbb{K}^n \otimes \mathbb{K}^m$, allora $R(M) = \underline{R}(M)$.*

Dimostrazione. Dall'algebra lineare sappiamo che una matrice M ha rango $\mathrm{rk}(M) \leq r$ se e solo se è possibile scriverla come somma di r matrici di rango $\mathrm{rk} = 1$. É immediato verificare che una matrice ha rango $\mathrm{rk} = 1$ se e solo se ha rango $R = 1$ come tensore. Pertanto le due nozioni di rango (per matrici e per tensori di ordine 2) coincidono. Per dimostrare l'asserto è sufficiente osservare che l'insieme

$$X_r = \{M \in \mathcal{M}(n \times m, \mathbb{K}) \text{ t.c. } \mathrm{rk}(M) \leq r\}$$

delle matrici di rango $\leq r$ è una varietà proiettiva. Infatti, dall'algebra lineare, sappiamo che le equazioni che definiscono X_r sono i determinanti dei minori $(r+1) \times (r+1)$, pertanto $X_r = \sigma_r\left(\mathrm{Seg}(\mathbb{P}^n \times \mathbb{P}^m)\right)$, concludendo così la dimostrazione. □

2.5 Spazio Tangente

Definizione. Sia $X \subseteq \mathbb{P}^n$ una varietà proiettiva, sia $[P] \in X$ un punto. Lo spazio tangente a X nel punto $[P]$ è l'insieme

$$T_{[P]}X = \left\{ [x_0, \ldots, x_n] \in \mathbb{P}^n \text{ t.c. } \sum_{i=0}^{n} \frac{\partial f}{\partial x_i}(P) \cdot x_i = 0 \text{ per ogni } f \in I(X) \right\}$$

Lemma 2.5.1. *(di Terracini) Siano $X_1, \ldots, X_k \subseteq \mathbb{P}^n$ varietà proiettive. Siano $[P_1] \in X_1, \ldots, [P_k] \in X_k, [P] \in J(X_1, \ldots, X_n)$ punti generali tali che $[P] \in \langle [P_1], \ldots, [P_k] \rangle$. Allora*

$$T_{[P]}J(X_1, \ldots, X_k) = \langle T_{[P_1]}X_1, \ldots, T_{[P_k]}X_k \rangle$$

Proposizione 2.5.1. *Sia $\mathbb{K}$ un campo e siano U, V spazi vettoriali di dimensione finita su $\mathbb{K}$ e sia $X = Seg(\mathbb{P}U \times \mathbb{P}V)$. Siano $u \in U, v \in V$ vettori non nulli, allora*

$$T_{[u \otimes v]}X = \mathbb{P}(U \otimes v + u \otimes V)$$

Dimostrazione. Fissiamo delle basi in U e in V. Per comodità possiamo supporre u e v come primi elementi di queste basi, dunque letti nelle basi $u = (1, 0, \ldots, 0)$ e $v = (1, 0, \ldots, 0)$. Lo spazio tangente $T_{[u \otimes v]}(X)$ è l'insieme delle matrici

$$x = \begin{bmatrix} x_{0,0} & \cdots & x_{0,b} \\ \vdots & \ddots & \vdots \\ x_{a,0} & \cdots & x_{a,b} \end{bmatrix}$$

per le quali $\sum_{i,j} \frac{\partial F}{\partial x_{i,j}}(u\otimes v)\cdot x_{i,j} = 0$ per ogni $F \in I(X)$. In particolare $I(X)$ è generato dai minori 2×2, pertanto considerando i polinomi della forma

$$F = x_{i_1,j_1}x_{i_2,j_2} - x_{i_1,j_2}x_{i_2,j_2}$$

per ogni $i_1 \neq i_2, j_1 \neq j_2$ abbiamo

$$\sum_{i,j} \frac{\partial F}{\partial x_{i,j}}(u \otimes v) \cdot x_{i,j} = u_{i_2}v_{j_2}x_{i_1,j_1} + u_{i_1}v_{j_1}x_{i_2,j_2} - u_{i_2}v_{j_1}x_{i_1,j_2} - u_{i_1}v_{j_2}x_{i_2,j_1}$$

dunque sostituendo i valori delle componenti di u e v: $u_i = \delta_{i,1}$ e $v_j = \delta_{j,1}$ abbiamo le equazioni

$$x_{i,j} = 0$$

per ogni $i \neq 1, j \neq 1$. Quindi possiamo concludere che lo spazio tangente è costituito dalle matrici della forma

$$x = \begin{bmatrix} x_{0,0} & x_{0,1} & \cdots & x_{0,b} \\ x_{1,0} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{a,0} & 0 & \cdots & 0 \end{bmatrix}$$

Ovvero sono esattamente gli elementi dello spazio $U \otimes v + u \otimes V$ a meno di un fattore moltiplicativo, e questo conclude la dimostrazione. □

Proposizione 2.5.2. *Sia $\mathbb{K}$ un campo e siano $V_1, \ldots, V_n$ spazi vettoriali di dimensione finita su $\mathbb{K}$. Sia $X = Seg(\mathbb{P}V_1 \times \ldots \times \mathbb{P}V_n)$ e siano $v_1 \in V_1, \ldots, v_n \in V_n$ vettori non nulli. Ponendo*

$$T_i = v_1 \otimes \ldots \otimes v_{i-1} \otimes V_i \otimes v_{i+1} \otimes \ldots \otimes v_n$$

abbiamo

$$T_{[v_1\otimes\ldots\otimes v_n]}X = \mathbb{P}\left(T_1 + \ldots + T_n\right)$$

Dimostrazione. La dimostrazione è analoga a quella della Proposizione 2.5.1. □

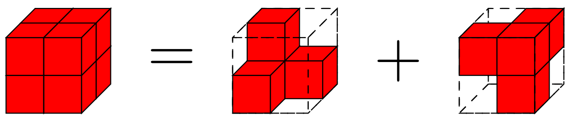
Proposizione 2.5.3. *Sia $X = Seg(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1)$, allora $\sigma_2(X) = \mathbb{P}^7$. Ovvero un tensore generale T di tipo $2 \times 2 \times 2$ ha rango bordo $\underline{R}(T) \leq 2$.*

Dimostrazione. Usando la proposizione 2.5.2 possiamo calcolare lo spazio tangente a $\sigma_2(X)$ nel punto generale $[P]$, con $P = u_1 \otimes v_1 \otimes w_1 + u_2 \otimes v_2 \otimes w_2$.

Possiamo assumere $\mathbb{C}^2 = \langle u_1, u_2 \rangle = \langle v_1, v_2 \rangle = \langle w_1, w_2 \rangle$, dunque

$$\begin{aligned} T_1 + T_2 &= \mathbb{C}^2 \otimes v_1 \otimes w_1 + u_1 \otimes \mathbb{C}^2 \otimes w_1 + u_1 \otimes v_1 \otimes \mathbb{C}^2 \\ &+ \mathbb{C}^2 \otimes v_2 \otimes w_2 + u_2 \otimes \mathbb{C}^2 \otimes w_2 + u_2 \otimes v_2 \otimes \mathbb{C}^2 \\ &= \mathbb{C}^8 \end{aligned}$$

Dunque $T_{[P]}\left(\sigma_2(X)\right) = \mathbb{P}^7$ e l'asserto è provato. $\square$



Capitolo 3

Prodotto di Matrici

Il rango fornisca una buona stima della complessità computazionale di un tensore. In [1] viene provato come la complessità asintotica del prodotto di matrici si possa stimare attraverso il rango. Vedremo in questo capitolo come ottenere informazioni utili sui possibili algoritmi minimali per un tensore assegnato, in particolare per il tensore prodotto di matrici. Come ultimo argomento vedremo una dimostrazione alternativa del teorema, provato indipendentemente in [8] e in [15], che per moltiplicare matrici 2×2 sono necessari almeno sette prodotti. Inoltre rimandiamo il lettore a [11] per un risultato ancora più forte, ovvero che anche il rango bordo per il prodotto di matrici 2×2 è sette.

Definizione. Sia $\mathbb{K}$ un campo e siano U, V, W spazi vettoriali su $\mathbb{K}$. Siano

$$\begin{aligned} A &= \mathrm{Hom}(U, V) \simeq U^* \otimes V \\ B &= \mathrm{Hom}(V, W) \simeq V^* \otimes W \\ C &= \mathrm{Hom}(U, W) \simeq U^* \otimes W \end{aligned}$$

definiamo il tensore *composizione di funzioni* l'applicazione $\Psi : A \times B \longrightarrow C$ ponendo:

$$\Psi(f, g) = g \circ f$$

Utilizzando isomorfismi naturali sui prodotti tensoriali è utile considerare Ψ come un tensore nei seguenti spazi

$$\begin{aligned} A^* \otimes B^* \otimes C &\simeq (U^* \otimes V)^* \otimes (V^* \otimes W)^* \otimes (U^* \otimes W) \\ &\simeq (U \otimes V^*) \otimes (V \otimes W^*) \otimes (U^* \otimes W) \\ &\simeq (U \otimes V \otimes W)^* \otimes (U \otimes V \otimes W) \\ &\simeq \mathrm{Hom}(U \otimes V \otimes W, U \otimes V \otimes W) \end{aligned}$$

Ovvero come

$$\Psi : (U^* \otimes V) \times (V^* \otimes W) \times (W^* \otimes U) \longrightarrow \mathbb{K}$$

definita ponendo

$$\Psi(\alpha \otimes v, \beta \otimes w, \gamma \otimes u) = \alpha(u)\beta(v)\gamma(w)$$

ed estesa linearmente.

In quanto segue poniamo $\dim(U) = m, \dim(V) = n, \dim(W) = m$ e per $a, b \in \mathbb{N}$ identifichiamo lo spazio di matrici $\mathcal{M}(a \times b, \mathbb{K})$ con $\mathrm{Hom}(\mathbb{K}^b, \mathbb{K}^a)$.

Definizione. Siano $l, n, m \in \mathbb{N}$. Il tensore *prodotto di matrici* $m \times n \times l$ è l'applicazione $\Psi_{\langle m,n,l\rangle} : \mathcal{M}(n \times l, \mathbb{K}) \times \mathcal{M}(m \times n, \mathbb{K}) \longrightarrow \mathcal{M}(m \times l, \mathbb{K})$ definita da

$$\Psi_{\langle m,n,l\rangle}(Y, X) = XY$$

Possiamo vedere $\Psi_{\langle m,n,l\rangle} \in \mathcal{M}(m \times n, \mathbb{K})^* \otimes \mathcal{M}(n \times l, \mathbb{K})^* \otimes \mathcal{M}(m \times l, \mathbb{K})$

Assegnate delle basi in A^*, B^*, C il tensore $\Psi_{\langle m,n,l\rangle}$ corrisponde dunque al tensore Ψ letto nelle basi. Se $\{\alpha_{i,j}\}_{i,j}, \{\beta_{i,j}\}_{i,j}, \{c_{j,i}\}_{i,j}$ sono le basi canoniche rispettivamente su A^*, B^*, C vale la decomposizione

$$\Psi_{\langle m,n,l\rangle} = \sum_{i,j,k} \alpha_{i,j} \otimes \beta_{j,k} \otimes c_{k,i}$$

Equivalentemente possiamo pensare

$$\begin{aligned} \Psi_{\langle m,n,l\rangle} &= \sum_i \left(\sum_{j,k} \alpha_{i,j} \otimes \beta_{j,k} \otimes c_{k,i} \right) \\ &= \sum_i (XYZ)_{i,i} \\ &= \mathrm{Tr}(XYZ) \end{aligned}$$

per le matrici X, Y, Z corrispondenti alle coordinate rispetto alle basi scelte in A^*, B^*, C. Pertanto $\Psi_{\langle m,n,l\rangle}$ è invariante rispetto ai seguenti cambi di base in U, V, W:

$$\mathrm{Tr}(XYZ) = \mathrm{Tr}((H^{-1}XK)(K^{-1}YM)(M^{-1}ZH))$$

per ogni $H \in \mathrm{GL}(U), K \in \mathrm{GL}(V), M \in \mathrm{GL}(W)$.

L'invarianza del tensore prodotto di matrici sotto cambiamenti di base in U, V, W è un fatto non sorprendente: per effettuare la composizione di funzioni mediante prodotto di matrici utilizziamo lo stesso algoritmo, indipendentemente dalla base scelta.

$$\begin{tikzcd}
U \arrow[r, "Y"] \arrow[rr, bend left, "XY"] & V \arrow[r, "X"] & W \\
U \arrow[u, "H"] \arrow[r, "Y'"] \arrow[rr, bend right, "(XY)' = X'Y'"'] & V \arrow[u, "K"] \arrow[r, "X'"] & W \arrow[u, "M"]
\end{tikzcd}$$

3.1 Algoritmo di Strassen

Seguendo l'algoritmo di Strassen riportato nell'introduzione possiamo ricostruire una decomposizione del tensore $\Psi_{\langle 2,2,2\rangle}$:

$$\begin{aligned}\Psi_{\langle 2,2,2\rangle} &= (\alpha_{1,1}+\alpha_{2,2})\otimes(\beta_{1,1}+\beta_{2,2})\otimes(c_{1,1}+c_{2,2})\\ &+\alpha_{1,1}\otimes(\beta_{1,2}-\beta_{2,2})\otimes(c_{2,1}+c_{2,2})\\ &+\alpha_{2,2}\otimes(\beta_{2,1}-\beta_{1,1})\otimes(c_{1,2}+c_{1,1})\\ &+(\alpha_{2,1}+\alpha_{2,2})\otimes\beta_{1,1}\otimes(c_{1,2}-c_{2,2})\\ &+(\alpha_{1,2}+\alpha_{1,1})\otimes\beta_{2,2}\otimes(c_{2,1}-c_{1,1})\\ &+(-\alpha_{1,1}+\alpha_{2,1})\otimes(\beta_{1,1}+\beta_{1,2})\otimes c_{2,2}\\ &+(-\alpha_{2,2}+\alpha_{1,2})\otimes(\beta_{2,2}+\beta_{2,1})\otimes c_{1,1}\end{aligned}$$

3.2 Varietà di Algoritmi Ottimali

L'algoritmo di Strassen non è l'unico algoritmo per il prodotto di matrici dotato di soli 7 addendi di rango 1. Infatti è possibile applicare un qualsiasi cambiamento di base negli spazi U, V, W all'algoritmo di Strassen per ottenerne uno equivalente. Inoltre, come provato in [3] e [4] ogni algoritmo ottimale per il prodotto di matrici 2×2 è equivalente all'algoritmo di Strassen mediante un opportuno cambiamento di basi. Vediamo brevemente il teorema senza entrare nelle questioni più tecniche.

Definizione. Sia $T \in A^* \otimes B^* \otimes C$. Un *algoritmo di lunghezza R per T* è una R-upla

$$(\alpha_1 \otimes \beta_1 \otimes c_1, \dots, \alpha_R \otimes \beta_R \otimes c_R)$$

tale che

$$T = \sum_{r=1}^{R} \alpha_r \otimes \beta_r \otimes c_r$$

Gli algoritmi di lunghezza $R(T)$ per T si dicono *ottimali.*

Osservazione. Sia $R = R(T)$ e sia $(\alpha_1\otimes\beta_1\otimes c_1, \dots, \alpha_R\otimes\beta_R\otimes c_R)$ un algoritmo ottimale per T. Sia $\{e_k\}_{k=1,\dots,d}$ una base per C. Possiamo scrivere

$$c_i = \sum_{k=1}^{d} \gamma_{i,k} e_k$$

per qualche $\gamma_{i,k} \in \mathbb{K}$, dunque

$$T = \sum_{k=1}^{d} \left(\sum_{r=1}^{R} \gamma_{i,k} (\alpha_r \otimes \beta_r) \right)$$

Quindi ponendo $T_k = \sum_{r=1}^{R} \gamma_{i,k}(\alpha_r \otimes \beta_r)$ abbiamo

$$\langle T_1, \ldots, T_d \rangle \subseteq \langle \alpha_1 \otimes \beta_1, \ldots, \alpha_R \otimes \beta_R \rangle$$

Inoltre i prodotti $\alpha_r \otimes \beta_r$ determinano univocamente i coefficienti $\gamma_{i,k}$

$$\sum_{k=1}^{d} \sum_{r=1}^{R} \gamma_{r,k}(\alpha_r \otimes \beta_r \otimes e_k) = \sum_{k=1}^{d} \sum_{r=1}^{R} \gamma'_{r,k}(\alpha_r \otimes \beta_r \otimes e_k)$$

Dunque

$$\sum_{k=1}^{d} \sum_{r=1}^{R} (\gamma_{r,k} - \gamma'_{r,k})(\alpha_r \otimes \beta_r \otimes e_k) = 0$$

Pertanto deve essere

$$\sum_{r=1}^{R} \gamma_{r,k}(\alpha_r \otimes \beta_r)$$

per ogni $k = 1, \ldots, d$. Per ottimalità dell'algoritmo i vettori $\alpha_1 \otimes \beta_1, \ldots, \alpha_R \otimes \beta_R$ devono essere linearmente indipendenti (altrimenti esistono algoritmi di lunghezza $R-1$ per T, contro l'ottimalità dell'algoritmo scelto), dunque $\gamma_{r,k} = \gamma'_{r,k}$ per ogni $r = 1, \ldots, R$ e ogni $k = 1, \ldots, d$.

Dunque possiamo descrivere l'algoritmo ottimale assegnato usando i prodotti $(\alpha_1 \otimes \beta_1, \ldots, \alpha_R \otimes \beta_R)$. Più precisamente assegnare un algoritmo costituito dai termini $\alpha_r \otimes \beta_r \otimes c_r$ non determina univocamente $\alpha_r \otimes \beta_r$ tuttavia sono univocamente determinati gli elementi $[\alpha_r \otimes \beta_r]$.

Definizione. Sia $T \in A^* \otimes B^* \otimes C$. La *varietà degli algoritmi ottimali per* T è la varietà definita dalle equazioni (omogenee)

$$\langle [T_1], \ldots, [T_d] \rangle \subseteq \langle [\alpha_1 \otimes \beta_1], \ldots, [\alpha_R \otimes \beta_R] \rangle$$

dove $T_1, \ldots, T_d$ sono le slice di T lungo C. Gli elementi $P_r = \alpha_r \otimes \beta_r$ e i corrispondenti $[P_r]$ sono chiamati *prodotti per* T.

Definizione. Il gruppo $\mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C)$ agisce su $A^* \otimes B^* \otimes C$ come segue

$$(\phi, \psi, \tau) : \alpha \otimes \beta \otimes c \longmapsto (\alpha \circ \phi^{-1}) \otimes (\beta \circ \psi^{-1}) \otimes \tau(c)$$

Il *gruppo di isotropia* di un tensore $T \in A^* \otimes B^* \otimes C$ è il gruppo

$$\Gamma_T = \{g \in \mathrm{GL}(A) \times \mathrm{GL}(B) \times \mathrm{GL}(C) \text{ t.c. } gT = T\}$$

Definiamo inoltre il gruppo

$$\Gamma^0 = \Gamma^0_{\Psi_{\langle 2,2,2\rangle}} \{g \in \Gamma_{\Psi_{\langle 2,2,2\rangle}} \text{ t.c. } g \text{ agisce per cambi di base in } U, V, W\}$$

Definizione. Sia $T \in A^* \otimes B^* \otimes C$. Un insieme di prodotti $\{P_1, \dots, P_q\}$ è *indipendente su* T se $\{T_1, \dots, T_d, P_1, \dots, P_q\}$ sono linearmente indipendenti.

Definizione. Sia $T \in A^* \otimes B^* \otimes C$. Un insieme di prodotti $\{P_1, \dots, P_q\}$ *genera un algoritmo per* T se

$$\{T_1, \dots, T_d\} \subseteq \langle \mathrm{Seg}(\mathbb{P}A^* \times \mathbb{P}B^*) \cap \langle P_1, \dots, P_q, T_1, \dots, T_d \rangle \rangle$$

Teorema 3.2.1. *1.* $G^0_{\Psi_{\langle 2,2,2\rangle}} = \Gamma^0_{\Psi_{\langle 2,2,2\rangle}} \mathfrak{S}_7$ *agisce transitivamente sulla varietà degli algoritmi ottimali per* $\Psi_{\langle 2,2,2\rangle}$

2. La varietà di algoritmi ottimali per $\Psi_{\langle 2,2,2\rangle}$ *ha dimensione nove.*

Dimostrazione. (Cenni) Il gruppo Γ^0 agisce sui tensori decomponibili. Possiamo rappresentare i fattori $\alpha, \beta \in \mathbb{C}^4$ come matrici 2×2 nel modo seguente:

$$x = (x_1, x_2, x_3, x_4) \longmapsto X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$$

Chiamiamo $(\mathrm{rk}(\alpha), \mathrm{rk}(\beta))$ il *tipo del prodotto* $\alpha \otimes \beta$. Nel nostro caso si possono presentare solo i tipi $(2,2), (2,1), (1,2), (1,1)$. Se $\{P_1, \dots, P_q\} \subseteq A^* \otimes B^*$ la *distribuzione dei tipi di* $\{P_1, \dots, P_q\}$ è il vettore

$$(m_{2,2}, m_{2,1}, m_{1,2}, m_{1,1})$$

dove $m_{i,j}$ è il numero di prodotti di tipo (i,j) in $\{P_1, \dots, P_q\}$.

Lemma 3.2.1. *Se* $m_{1,1} \geq 5$ *ci sono almeno tre prodotti di tipo* $(1,1)$ *indipendenti su* $\Psi_{\langle 2,2,2\rangle}$. *Se* $m_{1,1} \leq 4$ *ogni terna di prodotti di tipo diverso da* $(1,1)$ *è indipendente su* $\Psi_{\langle 2,2,2\rangle}$.

Nel caso $m_{1,1} \leq 4$ consideriamo siffatti tre prodotti P_1, P_2, P_3 di tipo diverso da $(1,1)$ indipendenti su $\Psi_{\langle 2,2,2\rangle}$. La distribuzione dei tipi di $\{P_1, P_2, P_3\}$ è una delle seguenti:

1. $(3,0,0,0)$
2. $(2,1,0,0)$ oppure $(2,0,1,0)$
3. $(1,2,0,0)$ oppure $(1,0,2,0)$
4. $(1,1,1,0)$
5. $(0,3,0,0)$ oppure $(0,0,3,0)$

6. $(0, 2, 1, 0)$ oppure $(0, 1, 2, 0)$

Proposizione 3.2.1. *Prodotti P_1, P_2, P_3 di tipo $(2, 2), (2, \gamma), (2, \delta)$ (con $\gamma, \delta \in \{1, 2\}$) rispettivamente non possono generare algoritmi per $\Psi_{\langle 2,2,2 \rangle}$.*

Possiamo pertanto escludere i casi $(1), (2), (3)$.

Proposizione 3.2.2. *Prodotti P_1, P_2, P_3 di tipo $(2, 1)$ non possono generare algoritmi per $\Psi_{\langle 2,2,2 \rangle}$.*

Possiamo escludere il caso (5).

Proposizione 3.2.3. *Prodotti P_1, P_2, P_3 di tipo $(2, 1), (1, 2), (2, \gamma)$ (con $\gamma \in \{1, 2\}$) rispettivamente non possono generare algoritmi per $\Psi_{\langle 2,2,2 \rangle}$.*

Possiamo escludere i casi (4) e (6). Dunque deve essere $m_{1,1} \geq 5$, pertanto possiamo scegliere tre prodotti P_1, P_2, P_3 di tipo $(1, 1)$ indipendenti su $\Psi_{\langle 2,2,2 \rangle}$.

Proposizione 3.2.4. *Ogni algoritmo ottimale per $\Psi_{\langle 2,2,2 \rangle}$ generato da prodotti P_1, P_2, P_3 di tipo $(1, 1)$ è $G^0_{\Psi_{\langle 2,2,2 \rangle}}$-equivalente all'algoritmo di Strassen.*

Questo prova la prima parte del teorema. Per provare la seconda consideriamo i prodotti $\alpha_i \otimes \beta_i \otimes c_i$ dell'algoritmo di Strassen. Abbiamo

$$\gamma = (\alpha_1 \otimes \beta_1 \otimes c_1, \ldots, \alpha_7 \otimes \beta_7 \otimes c_7)$$

Il gruppo $G^0 = \Gamma^0 \cdot \mathfrak{S}_7$ agisce transitivamente sulla varietà di algoritmi $\mathcal{V}$, e Γ^0 ha dimensione nove. Dunque $\mathcal{V}$ ha dimensione al più nove. Sia

$$G^0_\gamma = \{\phi \in G^0 \text{ t.c. } \phi\gamma = \gamma\}$$

Sia Γ^0_γ l'insieme dei $\phi \in \Gamma^0$ per cui $\phi\gamma = \pi\gamma$ per qualche permutazione $\pi \in \mathfrak{S}_7$. Abbiamo $G^0_\gamma \subseteq \Gamma^0_\gamma \cdot \mathfrak{S}_7$. Se proviamo che Γ^0_γ è un gruppo finito anche G^0_γ è finito. Osserviamo che il prodotto $\alpha_1 \otimes \beta_1$ è di tipo $(2, 2)$, dunque dato che il tipo è Γ^0_γ-invariante ogni elemento $\phi \in \Gamma^0_\gamma$ deve fissare $\alpha_1 \otimes \beta_1 \otimes c_1$ e permutare gli altri prodotti. È facile osservare che se $\phi(\alpha_2 \otimes \beta_2 \otimes c_2) = \alpha_2 \otimes \beta_2 \otimes c_2$ allora $\phi = 1$. Se $\phi(\alpha_2 \otimes \beta_2 \otimes c_2) = \alpha_3 \otimes \beta_3 \otimes c_3$ allora ϕ agisce su γ come la permutazione $(23)(46)(57)$. Se $\phi(\alpha_2 \otimes \beta_2 \otimes c_2) = \alpha_4 \otimes \beta_4 \otimes c_4$ allora ϕ si comporta come $(245)(376)$. Dunque Γ^0_γ è un gruppo di 6 elementi, pertanto G^0_γ è finito e anche G^0/G^0_γ ha dimensione nove. Dal momento che $\mathcal{V}$ è isomorfa a G^0/G^0_γ, anche $\mathcal{V}$ ha dimensione nove. Questo conclude la dimostrazione.

□

3.3 Il Tensore $\Psi_{\langle 2,2,2\rangle}$ ha rango 7

Lemma 3.3.1. *Sia* $\Psi_{\langle n,n,n\rangle} = \sum_{r=1}^{R} \alpha_r \otimes \beta_r \otimes c_r$ *tale che* $rk(\alpha_r) = rk(\beta_r) = rk(c_r) = 1$ *per ogni* r*, allora* $R \geq n^3$.

Dimostrazione. Consideriamo l'isomorfismo naturale

$$F : (U \otimes V^*) \otimes (V \otimes W^*) \otimes (U^* \otimes W) \longrightarrow (U \otimes V \otimes W)^* \otimes (U \otimes V \otimes W)$$

definito da

$$F : (u \otimes \beta) \otimes (v \otimes \gamma) \otimes (\alpha \otimes w) \longmapsto (\alpha \otimes \beta \otimes \gamma) \otimes (u \otimes v \otimes w)$$

Abbiamo

$$\begin{aligned} F(\Psi_{\langle n,n,n\rangle}) &= F\left(\sum_{r=1}^{R} \alpha_r \otimes \beta_r \otimes c_r\right) \\ &= \sum_{r=1}^{R} F(\alpha_r \otimes \beta_r \otimes c_r) \end{aligned}$$

La matrice associata a $F(\Psi_{\langle n,n,n\rangle})$ è Id_{n^3}, dunque

$$\begin{aligned} n^3 &= \mathrm{rk}\left(F(\Psi_{\langle n,n,n\rangle})\right) \\ &= \mathrm{rk}\left(\sum_{r=1}^{R} F(\alpha_r \otimes \beta_r \otimes c_r)\right) \\ &\leq \sum_{r=1}^{R} \mathrm{rk}(F(\alpha_r \otimes \beta_r \otimes c_r)) \end{aligned}$$

Poiché $\mathrm{rk}(\alpha_r) = \mathrm{rk}(\beta_r) = \mathrm{rk}(c_r) = 1$ abbiamo che $\mathrm{rk}(F(\alpha_r \otimes \beta_r \otimes c_r)) = 1$, dunque $n^3 \leq R$. □

Corollario 3.3.1. *Un algoritmo ottimale per* $\Psi_{\langle 2,2,2\rangle}$ *contiene almeno un addendo* $\alpha \otimes \beta \otimes c$ *tale che* $(rk(\alpha), rk(\beta), rk(c)) \neq (1,1,1)$

Dimostrazione. Consideriamo una decomposizione di $\Psi_{\langle 2,2,2\rangle}$ in R addendi. Se i ranghi degli addendi sono tutti $(1,1,1)$ allora $R \geq 8$. Dunque la decomposizione non è ottimale (l'algoritmo di Strassen fornisce una decomposizione migliore). Dunque esiste almeno un addendo con ranghi distinti da $(1,1,1)$. □

Siano U, V, W spazi vettoriali. Dato un tensore $T \in U \otimes V \otimes W$ possiamo costruire un particolare flattening.

3.4 Flattening di Koszul

Definizione. Sia $T \in U \otimes V \otimes W$, il *flattening di Koszul* di T è l'applicazione lineare $T_U^\Lambda : U \otimes V^* \longrightarrow \Lambda^2 U \otimes W$ definita considerando il flattening $T_V : V^* \longrightarrow U \otimes W$ e prendendo la parte alternante di $T_V \otimes Id_U$ in $\Lambda^2 U \otimes W$.

Proposizione 3.4.1. *Sia $\mathbb{K}$ un campo e siano U, V, W spazi vettoriali di dimensione finita su $\mathbb{K}$. Siano $T, F \in U \otimes V \otimes W$, allora $(T+F)_U^\Lambda = T_U^\Lambda + F_U^\Lambda$.*

Teorema 3.4.1. *Siano A, B, C spazi vettoriali su $\mathbb{K}$ di dimensioni a, b, c. Sia $T \in A \otimes B \otimes C$ tale che $R(T) \leq R$. Allora $rk(T_A^\Lambda) \leq R\binom{a-1}{2}$.*

Dimostrazione. Possiamo scrivere $T = \sum_{r=1}^R a_r \otimes b_r \otimes c_r$, dunque:

$$\begin{aligned}
T_A^\Lambda &= \sum_{r=1}^R (a_r \otimes b_r \otimes c_r)_A^\Lambda \\
rk(T_A^\Lambda) &\leq \sum_{r=1}^R rk((a_r \otimes b_r \otimes c_r)_A^\Lambda) \\
&\leq \sum_{r=1}^R dim((a_r \wedge \Lambda^2(A)) \otimes c_r) \\
&\leq \sum_{r=1}^R \binom{a-1}{2} \\
&\leq R\binom{a-1}{2}
\end{aligned}$$

□

Corollario 3.4.1. *Se $rk(T_A^\Lambda) > R\binom{a-1}{2}$ allora $R(T) > r$.*

Corollario 3.4.2. *Nel caso $a = b = c = 4$ e $p = 2$ se $t \in A \otimes B \otimes C$ e $rk(T_A^\Lambda) \geq 16$ allora $R(T) \geq 6$.*

Attraverso l'azione del gruppo $GL(U) \times GL(V) \times GL(W)$ possiamo ottenere forme canoniche per tensori di rango piccolo. Queste saranno utili in seguito per effettuare calcoli espliciti.

Lemma 3.4.1. *Siano $Y, Z \in \mathcal{M}_2(\mathbb{C})$ tali che $rk(Y) = rk(Z) = 1$. Se lo spazio righe di Y non è ortogonale allo spazio colonne di Z allora esistono $H, K, M \in GL_2(\mathbb{C})$ tali che:*

$$K^{-1}YM = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$M^{-1}ZH = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Dimostrazione. Cerchiamo H, K, M invertibili tali che:

$$YM = K \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} k_{1,1} & 0 \\ k_{2,1} & 0 \end{bmatrix}$$

$$M^{-1}Z = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} H^{-1} = \frac{1}{det(H)} \begin{bmatrix} h_{2,2} & -h_{1,2} \\ 0 & 0 \end{bmatrix}$$

Supponiamo $\left[\begin{smallmatrix} y_{1,1} & y_{1,2} \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 & 0 \end{smallmatrix}\right]$ e $\left[\begin{smallmatrix} z_{1,1} \\ z_{2,1} \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ a meno di scegliere l'altra riga di Y e l'altra colonna di Z. Poniamo

$$M = \begin{bmatrix} z_{1,1} & y_{1,2} \\ z_{2,1} & -y_{1,1} \end{bmatrix}$$

Se M non fosse invertibile sarebbe $rk(M) = 1$, quindi $\left[\begin{smallmatrix} z_{1,1} \\ z_{2,1} \end{smallmatrix}\right]$ sarebbe parallelo a $\left[\begin{smallmatrix} y_{1,2} \\ -y_{1,1} \end{smallmatrix}\right]$. Tuttavia $\left[\begin{smallmatrix} y_{1,2} \\ -y_{1,1} \end{smallmatrix}\right]$ è ortogonale a $\left[\begin{smallmatrix} y_{1,1} \\ y_{1,2} \end{smallmatrix}\right]$, quindi in tal caso anche $\left[\begin{smallmatrix} y_{1,1} \\ y_{1,2} \end{smallmatrix}\right]$ sarebbe ortogonale a $\left[\begin{smallmatrix} z_{1,1} \\ z_{2,1} \end{smallmatrix}\right]$, dunque abbiamo un assurdo e M è invertibile. Rimangono da determinare H e K. Sostituendo

$$\begin{aligned} YM &= Y \begin{bmatrix} z_{1,1} & y_{1,2} \\ z_{2,1} & -y_{1,1} \end{bmatrix} \\ &= \begin{bmatrix} y_{1,1}z_{1,1} + y_{1,2}z_{2,1} & y_{1,1}y_{1,2} - y_{1,2}y_{1,1} \\ y_{2,1}z_{1,1} + y_{2,2}z_{2,1} & y_{2,1}y_{1,2} - y_{2,2}y_{1,1} \end{bmatrix} \\ &= \begin{bmatrix} y_{1,1}z_{1,1} + y_{1,2}z_{2,1} & 0 \\ y_{2,1}z_{1,1} + y_{2,2}z_{2,1} & -det(Y) \end{bmatrix} \\ &= \begin{bmatrix} y_{1,1}z_{1,1} + y_{1,2}z_{2,1} & 0 \\ y_{2,1}z_{1,1} + y_{2,2}z_{2,1} & 0 \end{bmatrix} \end{aligned}$$

Dunque deve essere

$$\begin{bmatrix} y_{1,1}z_{1,1} + y_{1,2}z_{2,1} & 0 \\ y_{2,1}z_{1,1} + y_{2,2}z_{2,1} & 0 \end{bmatrix} = \begin{bmatrix} k_{1,1} & 0 \\ k_{2,1} & 0 \end{bmatrix}$$

Dal momento che $y_{1,1}z_{1,1} + y_{1,2}z_{2,1} \neq 0$ (per ipotesi) possiamo porre $k_{1,1} = y_{1,1}z_{1,1} + y_{1,2}z_{2,1}$ e completare K a una matrice invertibile. Analogamente sostituendo

$$\begin{aligned} M^{-1}Z &= \frac{1}{det(M)} \begin{bmatrix} -y_{1,1} & -y_{1,2} \\ -z_{2,1} & z_{1,1} \end{bmatrix} \\ &= \frac{1}{det(M)} \begin{bmatrix} -y_{1,1}z_{1,1} - y_{1,2}z_{2,1} & -y_{1,1}z_{1,2} - y_{1,2}z_{2,2} \\ -z_{2,1}z_{1,1} + z_{1,1}z_{2,1} & -z_{2,1}z_{1,2} + z_{1,1}z_{2,2} \end{bmatrix} \\ &= \frac{1}{det(M)} \begin{bmatrix} -y_{1,1}z_{1,1} - y_{1,2}z_{2,1} & -y_{1,1}z_{1,2} - y_{1,2}z_{2,2} \\ 0 & -det(Z) \end{bmatrix} \\ &= \frac{1}{det(M)} \begin{bmatrix} -y_{1,1}z_{1,1} - y_{1,2}z_{2,1} & -y_{1,1}z_{1,2} - y_{1,2}z_{2,2} \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Pertanto deve essere

$$\frac{1}{det(H)} \begin{bmatrix} h_{2,2} & -h_{1,2} \\ 0 & 0 \end{bmatrix} = \frac{1}{det(M)} \begin{bmatrix} -y_{1,1}z_{1,1} - y_{1,2}z_{2,1} & -y_{1,1}z_{1,2} - y_{1,2}z_{2,2} \\ 0 & 0 \end{bmatrix}$$

Essendo $-y_{1,1}z_{1,1} - y_{1,2}z_{2,1} \neq 0$ per ipotesi è possibile determinare H invertibile tale che

$$\begin{cases} h_{2,2} = \frac{\det(M)}{\det(H)}(-y_{1,1}z_{1,1} - y_{1,2}z_{2,1}) \\ -h_{1,2} = \frac{\det(M)}{\det(H)}(-y_{1,1}z_{1,2} - y_{1,2}z_{2,2}) \end{cases}$$

□

Lemma 3.4.2. *Siano $Y, Z \in \mathcal{M}_2(\mathbb{C})$ tali che $rk(Y) = rk(Z) = 1$. Se lo spazio righe di Y è ortogonale allo spazio colonne di Z allora esistono $H, K, M \in GL_2(\mathbb{C})$ tali che:*

$$K^{-1}YM = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$M^{-1}ZH = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Dimostrazione. Cerchiamo H, K, M invertibili tali che:

$$YM = K \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} k_{1,1} & 0 \\ k_{2,1} & 0 \end{bmatrix}$$

$$M^{-1}Z = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} H^{-1} = \frac{1}{det(H)} \begin{bmatrix} 0 & 0 \\ h_{2,2} & -h_{1,2} \end{bmatrix}$$

Supponiamo $\left[\begin{smallmatrix} z_{1,1} \\ z_{2,1} \end{smallmatrix}\right] \neq \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$ a meno di scegliere l'altra colonna di Z. Poniamo

$$M = \begin{bmatrix} -z_{2,1} & z_{1,1} \\ z_{1,1} & z_{2,1} \end{bmatrix}$$

Allora M è invertibile. Sostituendo

$$\begin{aligned} YM &= Y \begin{bmatrix} -z_{2,1} & z_{1,1} \\ z_{1,1} & z_{2,1} \end{bmatrix} \\ &= \begin{bmatrix} -y_{1,1}z_{2,1} + y_{1,2}z_{1,1} & y_{1,1}z_{1,1} + y_{1,2}z_{2,1} \\ -y_{2,1}z_{2,1} + y_{2,1}z_{1,1} & y_{2,1}z_{1,1} + y_{2,2}z_{2,1} \end{bmatrix} \\ &= \begin{bmatrix} -y_{1,1}z_{2,1} + y_{1,2}z_{1,1} & 0 \\ -y_{2,1}z_{2,1} + y_{2,1}z_{1,1} & 0 \end{bmatrix} \end{aligned}$$

Dunque ponendo

$$\begin{cases} k_{1,1} = -y_{1,1}z_{2,1} + y_{1,2}z_{1,1} \\ k_{2,1} = -y_{2,1}z_{2,1} + y_{2,2}z_{1,1} \end{cases}$$

uno almeno tra $k_{1,1}$ e $k_{2,1}$ è diverso da 0, quindi possiamo estendere a K invertibile. Analogamente sostituendo

$$\begin{aligned} M^{-1}Z &= \frac{1}{det(M)} \begin{bmatrix} z_{2,1} & -z_{1,1} \\ -z_{1,1} & -z_{2,1} \end{bmatrix} Z \\ &= \frac{1}{det(M)} \begin{bmatrix} z_{2,1}z_{1,1} - z_{1,1}z_{2,1} & z_{2,1}z_{1,2} - z_{1,1}z_{2,2} \\ -z_{1,1}z_{1,1} - z_{2,1}z_{2,1} & -z_{1,1}z_{2,1} - z_{2,1}z_{2,2} \end{bmatrix} \\ &= \frac{1}{det(M)} \begin{bmatrix} 0 & -det(Z) \\ -z_{1,1}z_{1,1} - z_{2,1}z_{2,1} & -z_{1,1}z_{2,1} - z_{2,1}z_{2,2} \end{bmatrix} \\ &= \frac{1}{det(M)} \begin{bmatrix} 0 & 0 \\ -z_{1,1}z_{1,1} - z_{2,1}z_{2,1} & -z_{1,1}z_{2,1} - z_{2,1}z_{2,2} \end{bmatrix} \end{aligned}$$

Dunque deve essere

$$\frac{1}{det(M)} \begin{bmatrix} 0 & 0 \\ -z_{1,1}z_{1,1} - z_{2,1}z_{2,1} & -z_{1,1}z_{2,1} - z_{2,1}z_{2,2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ h_{2,2} & -h_{1,2} \end{bmatrix}$$

Essendo $-z_{1,1}z_{1,1} - z_{2,1}z_{2,1} \neq 0$ possiamo trovare H invertibile tale che

$$\begin{cases} h_{2,2} = \frac{det(H)}{det(M)}(-z_{1,1}z_{1,1} - z_{2,1}z_{2,1}) \\ -h_{1,2} = \frac{det(H)}{det(M)}(-z_{1,1}z_{2,1} - z_{2,1}z_{2,2}) \end{cases}$$

□

Teorema 3.4.2. $R(\Psi_{\langle 2,2,2\rangle}) = 7$

Dimostrazione. Proviamo che $R(\Psi_{\langle 2,2,2\rangle}) \geq 7$ (l'altra diseguaglianza è fornita dall'algoritmo di Strassen). Sia $r = R(t_{\langle 2,2,2\rangle})$, consideriamo una decomposizione di $\Psi_{\langle 2,2,2\rangle}$ in r addendi di rango 1:

$$\Psi_{\langle 2,2,2\rangle} = \sum_{i=1}^{r} a_i \otimes b_i \otimes c_i$$

Per il Lemma 3.3.1 tra questi addendi almeno uno, che indichiamo con $x \otimes y \otimes z$, ha ranghi diversi da $(1,1,1)$. Sia $F = \Psi_{\langle 2,2,2\rangle} - x \otimes y \otimes z$. Fissata una base esprimiamo x con X, y con Y, z con Z, è sufficiente provare che $R(F) \geq 6$.

1. Caso $(2,2,*)$

   Usando il cambio di base definito da $H = X$, $K = \mathrm{Id}_2$, $M = Y^{-1}$ possiamo porre $X = Y = \mathrm{Id}_2$. Sia $Z = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Calcolando F con Macaulay2 otteniamo

```
| 0    0    0  a    0 0 0 0  0  0 0 0 -1 0 0 0  0  0  0  0  0 a   0   0   |
| -a+1 0    0  0    a 0 0 0  0  0 0 0 0  0 0 0  0  0  -a 0  0 0   a   0   |
| 0    -a+1 0  0    0 a 0 0  0  0 0 0 0  0 1 0  0  0  0  -a 0 0   0   a   |
| 0    0    0  -a+1 0 0 0 0  0  0 0 0 0  0 0 0  1  0  0  0  0 -a  0   0   |
| 0    -1   0  b    0 0 0 0  0  0 0 0 0  0 0 -1 0  0  0  0  0 b   0   0   |
| -b   0    -1 0    b 0 0 0  0  0 0 0 0  0 0 0  -1 0  -b 0  0 0   b   0   |
| 0    -b   0  0    0 b 0 0  0  0 0 0 0  0 0 0  0  -1 0  -b 0 0   0   b   |
| 0    0    0  -b   0 1 0 0  0  0 0 0 0  0 0 0  0  0  0  0  0 -b  0   0   |
| 0    0    0  c    0 0 0 0  0  0 0 0 0  0 0 0  0  0  -1 0  0 c   0   0   |
| -c   0    0  0    c 0 1 0  0  0 0 0 0  0 0 0  0  0  -c 0  0 0   c   0   |
| 0    -c   0  0    0 c 0 1  0  0 0 0 0  0 0 0  0  0  0  -c 1 0   0   c   |
| 0    0    0  -c   0 0 0 0  0  1 0 0 0  0 0 0  0  0  0  0  0 -c  1   0   |
| 0    0    0  d    0 0 0 -1 0  0 0 0 0  0 0 0  0  0  0  0  0 d-1 0   0   |
| -d   0    0  0    d 0 0 0  -1 0 0 0 0  0 0 0  0  0  -d 0  0 0   d-1 0   |
| 0    -d   0  0    0 d 0 0  0  0 0 0 0  0 0 0  0  0  0  -d 0 0   0   d-1 |
| 0    0    0  -d   0 0 0 0  0  0 0 1 0  0 0 0  0  0  0  0  0 -d  0   0   |
```

   Distinguiamo 4 casi

   (a) $a = 1$ e $d = 1$

   Selezionando le colonne $\{11, 1, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 4, 12\}$ (in quest'ordine) otteniamo

```
| 0 0  0  0  0 0 0 0  0 0  0  0  0  0 0 -1 |
| 0 0  0  0  0 0 0 -1 0 0  0  0  0  0 1 0  |
| 0 0  0  0  0 0 0 0  1 0  0  0  0  1 0 0  |
| 0 0  0  0  0 0 0 0  0 0  0  0  1  0 0 0  |
```

```
| 0 -1 0  0  0 0 0 0  0 0  0  -1 0  0 0 0  |
| 0 0  0  0  0 0 0 -b 0 0  -1 0  -1 0 b 0  |
| 0 -b 0  0  0 0 0 0  b -1 0  0  0  0 0 0  |
| 0 0  0  0  0 0 0 0  1 0  0  0  0  0 0 0  |
| 0 0  0  0  0 0 0 -1 0 0  0  0  0  0 0 0  |
| 0 0  0  0  0 0 1 -c 0 0  0  0  0  0 c 0  |
| 0 -c 0  1  0 1 0 0  c 0  0  0  0  0 0 0  |
| 0 0  0  0  1 0 0 0  0 0  0  0  0  0 0 0  |
| 0 0  0  -1 0 0 0 0  0 0  0  0  0  0 0 0  |
| 0 0  -1 0  0 0 0 -1 0 0  0  0  0  0 1 0  |
| 0 -1 0  0  0 0 0 0  1 0  0  0  0  0 0 0  |
| 1 0  0  0  0 0 0 0  0 0  0  0  0  0 0 0  |
```

Questa matrice ha determinante 1, dunque è invertibile.

(b) $a = 1$ e $d \neq 1$

Selezionando le colonne $\{11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 4, 12\}$ (in quest'ordine) otteniamo

```
| 0 0   0  0  0 0 0 0  0 0  0  0  0  0 0 -1 |
| 0 0   0  0  0 0 0 -1 0 0  0  0  0  0 1 0  |
| 0 1   0  0  0 0 0 0  1 0  0  0  0  1 0 0  |
| 0 0   0  0  0 0 0 0  0 0  0  0  1  0 0 0  |
| 0 0   0  0  0 0 0 0  0 0  0  -1 0  0 0 0  |
| 0 0   0  0  0 0 0 -b 0 0  -1 0  -1 0 b 0  |
| 0 b   0  0  0 0 0 0  b -1 0  0  0  0 0 0  |
| 0 0   0  0  0 0 0 0  1 0  0  0  0  0 0 0  |
| 0 0   0  0  0 0 0 -1 0 0  0  0  0  0 0 0  |
| 0 0   0  0  0 0 1 -c 0 0  0  0  0  0 c 0  |
| 0 c   0  1  0 1 0 0  c 0  0  0  0  0 0 0  |
| 0 0   0  0  1 0 0 0  0 0  0  0  0  0 0 0  |
| 0 0   0  -1 0 0 0 0  0 0  0  0  0  0 0 0  |
| 0 0   -1 0  0 0 0 -d 0 0  0  0  0  0 d 0  |
| 0 d-1 0  0  0 0 0 0  d 0  0  0  0  0 0 0  |
| 1 0   0  0  0 0 0 0  0 0  0  0  0  0 0 0  |
```

Questa ha determinante $-d + 1 \neq 0$, dunque è invertibile.

(c) $a \neq 1$ e $d = 1$

Selezionando le colonne $\{11, 1, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 0, 12\}$ (in quest'ordine) otteniamo

```
| 0 0    0  0  0 0 0 0  0 0  0  0  0  0 0       -1 |
| 0 0    0  0  0 0 0 -a 0 0  0  0  0  0 -a+1 0     |
| 0 -a+1 0  0  0 0 0 0  a 0  0  0  0  1 0     0    |
| 0 0    0  0  0 0 0 0  0 0  0  0  1  0 0     0    |
| 0 -1   0  0  0 0 0 0  0 0  0  -1 0  0 0     0    |
| 0 0    0  0  0 0 0 -b 0 0  -1 0  -1 0 -b    0    |
| 0 -b   0  0  0 0 0 0  b -1 0  0  0  0 0     0    |
| 0 0    0  0  0 0 0 0  1 0  0  0  0  0 0     0    |
| 0 0    0  0  0 0 0 -1 0 0  0  0  0  0 0     0    |
| 0 0    0  0  0 0 1 -c 0 0  0  0  0  0 -c    0    |
| 0 -c   0  1  0 1 0 0  c 0  0  0  0  0 0     0    |
| 0 0    0  0  1 0 0 0  0 0  0  0  0  0 0     0    |
| 0 0    0  -1 0 0 0 0  0 0  0  0  0  0 0     0    |
| 0 0    -1 0  0 0 0 -1 0 0  0  0  0  0 -1    0    |
```

```
| 0 -1   0  0  0 0 0 0  1 0  0  0  0  0 0    0  |
| 1 0    0  0  0 0 0 0  0 0  0  0  0  0 0    0  |
```

Questa ha determinante $-a+1 \neq 0$, dunque è invertibile.

(d) $a \neq 1$ e $d \neq 1$

Selezionando le colonne $\{11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 0, 12\}$ (in quest'ordine) otteniamo

```
| 0 0    0  0  0 0 0 0  0 0  0  0  0  0 0    -1 |
| 0 0    0  0  0 0 0 -a 0 0  0  0  0  0 -a+1 0  |
| 0 a    0  0  0 0 0 0  a 0  0  0  0  1 0    0  |
| 0 0    0  0  0 0 0 0  0 0  0  0  1  0 0    0  |
| 0 0    0  0  0 0 0 0  0 0  0  -1 0  0 0    0  |
| 0 0    0  0  0 0 0 -b 0 0  -1 0  -1 0 -b   0  |
| 0 b    0  0  0 0 0 0  b -1 0  0  0  0 0    0  |
| 0 0    0  0  0 0 0 0  1 0  0  0  0  0 0    0  |
| 0 0    0  0  0 0 0 -1 0 0  0  0  0  0 0    0  |
| 0 0    0  0  0 0 1 -c 0 0  0  0  0  0 -c   0  |
| 0 c    0  1  0 1 0 0  c 0  0  0  0  0 0    0  |
| 0 0    0  0  1 0 0 0  0 0  0  0  0  0 0    0  |
| 0 0    0  -1 0 0 0 0  0 0  0  0  0  0 0    0  |
| 0 0    -1 0  0 0 0 -d 0 0  0  0  0  0 -d   0  |
| 0 d-1  0  0  0 0 0 0  d 0  0  0  0  0 0    0  |
| 1 0    0  0  0 0 0 0  0 0  0  0  0  0 0    0  |
```

Questa ha determinante $(a-1)(d-1) \neq 0$, dunque è invertibile.

Nei casi appena visti $\mathrm{rk}(F_A^{\Lambda p}) = 16$, dunque $R(F) \geq 6$.

2. Casi $(2, *, 2)$ e $(*, 2, 2)$

   Per simmetria, scambiando il ruolo di A, B, C.

3. Caso $(2, 1, 1)$

   Il Lemma 3.4.1 e il Lemma 3.4.2 ci consentono di studiare il problema in due soli casi. In entrambi poniamo $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

   (a) Primo caso

$$Y = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Calcolando con Macaulay2 otteniamo

```
| b    c    0  d    0  0  0 0 0 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 0 |
| -a+1 0    c  0    d  0  0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 |
| 0    -a+1 -b 0    0  d  0 0 0 0 0 0 0  0 1 0 0 0 0 0 0 0 0 0 |
| 0    0    0  -a+1 -b -c 0 0 0 0 0 0 0  0 0 0 1 0 0 0 0 0 0 0 |
```

```
| 0   -1   0  0    0  0  0 0   0   0 0 0 0   0 0 -1 0   0   0   0 0 0   0   0  |
| 0   0   -1  0    0  0  0 0   0   0 0 0 0   0 0 0   -1 0   0   0 0 0   0   0  |
| 0   0    0  0    0  0  0 0   0   0 0 0 0   0 0 0   0   -1 0   0 0 0   0   0  |
| 0   0    0  0    0  1  0 0   0   0 0 0 0   0 0 0   0   0   0   0 0 0   0   0  |
| 0   0    0  0    0  0  0 0   0   0 0 0 0   0 0 0   0   0  -1   0 0 0   0   0  |
| 0   0    0  0    0  0  1 0   0   0 0 0 0   0 0 0   0   0   0   0 0 0   0   0  |
| 0   0    0  0    0  0  0 1   0   0 0 0 0   0 0 0   0   0   0   0 1 0   0   0  |
| 0   0    0  0    0  0  0 0   0   1 0 0 0   0 0 0   0   0   0   0 0 0   1   0  |
| 0   0    0  0    0  0  0 -1  0   0 0 0 0   0 0 0   0   0   0   0 0 -1  0   0  |
| 0   0    0  0    0  0  0 0  -1   0 0 0 0   0 0 0   0   0   0   0 0 0  -1   0  |
| 0   0    0  0    0  0  0 0   0   0 0 0 0   0 0 0   0   0   0   0 0 0   0  -1  |
| 0   0    0  0    0  0  0 0   0   0 0 1 0   0 0 0   0   0   0   0 0 0   0   0  |
```

Distinguiamo due ulteriori casi

i. $a = 1$

Selezionando le colonne $\{11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 1, 16, 14, 4, 12\}$ (in quest'ordine) otteniamo

```
| 0 0  0  0   0 0 0 0   0  0  0  c     0  0 0  -1 |
| 0 0  0  0   0 0 0 0   0  0  c  0     0  0 d  0  |
| 0 0  0  0   0 0 0 0   d  0  -b -a+1  0  1 0  0  |
| 0 0  0  0   0 0 0 0   -c 0  0  0     1  0 -b 0  |
| 0 0  0  0   0 0 0 0   0  0  0  -1    0  0 0  0  |
| 0 0  0  0   0 0 0 0   0  0  -1 0     -1 0 0  0  |
| 0 0  0  0   0 0 0 0   0  -1 0  0     0  0 0  0  |
| 0 0  0  0   0 0 0 0   1  0  0  0     0  0 0  0  |
| 0 0  0  0   0 0 0 -1  0  0  0  0     0  0 0  0  |
| 0 0  0  0   0 0 1 0   0  0  0  0     0  0 0  0  |
| 0 0  0  1   0 1 0 0   0  0  0  0     0  0 0  0  |
| 0 0  0  0   1 0 0 0   0  0  0  0     0  0 0  0  |
| 0 0  0  -1  0 0 0 0   0  0  0  0     0  0 0  0  |
| 0 0  -1 0   0 0 0 0   0  0  0  0     0  0 0  0  |
| 0 -1 0  0   0 0 0 0   0  0  0  0     0  0 0  0  |
| 1 0  0  0   0 0 0 0   0  0  0  0     0  0 0  0  |
```

Questa matrice ha determinante $-bc+d = ad-bc = \det(X) \neq 0$, dunque è invertibile.

ii. $a \neq 1$

Selezionando le colonne $\{11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 1, 16, 14, 0, 12\}$ (in quest'ordine) otteniamo

```
| 0 0  0  0   0 0 0 0   0  0  0  c     0  0 b     -1 |
| 0 0  0  0   0 0 0 0   0  0  c  0     0  0 -a+1  0  |
| 0 0  0  0   0 0 0 0   d  0  -b -a+1  0  1 0     0  |
| 0 0  0  0   0 0 0 0   -c 0  0  0     1  0 0     0  |
| 0 0  0  0   0 0 0 0   0  0  0  -1    0  0 0     0  |
| 0 0  0  0   0 0 0 0   0  0  -1 0     -1 0 0     0  |
| 0 0  0  0   0 0 0 0   0  -1 0  0     0  0 0     0  |
| 0 0  0  0   0 0 0 0   1  0  0  0     0  0 0     0  |
| 0 0  0  0   0 0 0 -1  0  0  0  0     0  0 0     0  |
| 0 0  0  0   0 0 1 0   0  0  0  0     0  0 0     0  |
| 0 0  0  1   0 1 0 0   0  0  0  0     0  0 0     0  |
| 0 0  0  0   1 0 0 0   0  0  0  0     0  0 0     0  |
| 0 0  0  -1  0 0 0 0   0  0  0  0     0  0 0     0  |
| 0 0  -1 0   0 0 0 0   0  0  0  0     0  0 0     0  |
```

```
| 0 -1 0  0  0 0 0 0  0  0  0  0    0  0 0    0  |
| 1 0  0  0  0 0 0 0  0  0  0  0    0  0 0    0  |
```

Questa ha determinante $-a+1 \neq 0$, dunque è invertibile.

(b) Secondo caso

$$Y = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$Z = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

In questo caso otteniamo

```
| 0  0  0  0  0  0  0 0  0  0 0 0 -1 0 0 0  0  0  0  0 0 0  0  0  |
| 1  0  0  0  0  0  0 0  0  0 0 0 0  0 0 0  0  0  0  0 0 0  0  0  |
| 0  1  0  0  0  0  0 0  0  0 0 0 0  0 1 0  0  0  0  0 0 0  0  0  |
| 0  0  0  1  0  0  0 0  0  0 0 0 0  0 0 0  1  0  0  0 0 0  0  0  |
| 0  -1 0  0  0  0  0 0  0  0 0 0 0  0 0 -1 0  0  0  0 0 0  0  0  |
| 0  0  -1 0  0  0  0 0  0  0 0 0 0  0 0 0  -1 0  0  0 0 0  0  0  |
| 0  0  0  0  0  0  0 0  0  0 0 0 0  0 0 0  0  -1 0  0 0 0  0  0  |
| 0  0  0  0  0  1  0 0  0  0 0 0 0  0 0 0  0  0  0  0 0 0  0  0  |
| b  c  0  d  0  0  0 0  0  0 0 0 0  0 0 0  0  0  -1 0 0 0  0  0  |
| -a 0  c  0  d  0  1 0  0  0 0 0 0  0 0 0  0  0  0  0 0 0  0  0  |
| 0  -a -b 0  0  d  0 1  0  0 0 0 0  0 0 0  0  0  0  0 1 0  0  0  |
| 0  0  0  -a -b -c 0 0  0  1 0 0 0  0 0 0  0  0  0  0 0 0  1  0  |
| 0  0  0  0  0  0  0 -1 0  0 0 0 0  0 0 0  0  0  0  0 0 -1 0  0  |
| 0  0  0  0  0  0  0 0  -1 0 0 0 0  0 0 0  0  0  0  0 0 0  -1 0  |
| 0  0  0  0  0  0  0 0  0  0 0 0 0  0 0 0  0  0  0  0 0 0  0  -1 |
| 0  0  0  0  0  0  0 0  0  0 0 1 0  0 0 0  0  0  0  0 0 0  0  0  |
```

Selezionando le colonne $\{12, 0, 1, 3, 15, 2, 17, 5, 18, 6, 7, 9, 21, 8, 23, 11\}$ (in quest'ordine) otteniamo la sottomatrice

```
| -1 0  0  0  0  0  0  0  0  0 0  0 0  0  0  0 |
| 0  1  0  0  0  0  0  0  0  0 0  0 0  0  0  0 |
| 0  0  1  0  0  0  0  0  0  0 0  0 0  0  0  0 |
| 0  0  0  1  0  0  0  0  0  0 0  0 0  0  0  0 |
| 0  0  -1 0  -1 0  0  0  0  0 0  0 0  0  0  0 |
| 0  0  0  0  0  -1 0  0  0  0 0  0 0  0  0  0 |
| 0  0  0  0  0  0  -1 0  0  0 0  0 0  0  0  0 |
| 0  0  0  0  0  0  0  1  0  0 0  0 0  0  0  0 |
| 0  b  c  d  0  0  0  0  -1 0 0  0 0  0  0  0 |
| 0  -a 0  0  0  c  0  0  0  1 0  0 0  0  0  0 |
| 0  0  -a 0  0  -b 0  d  0  0 1  0 0  0  0  0 |
| 0  0  0  -a 0  0  0  -c 0  0 0  1 0  0  0  0 |
| 0  0  0  0  0  0  0  0  0  0 -1 0 -1 0  0  0 |
| 0  0  0  0  0  0  0  0  0  0 0  0 0  -1 0  0 |
| 0  0  0  0  0  0  0  0  0  0 0  0 0  0  -1 0 |
| 0  0  0  0  0  0  0  0  0  0 0  0 0  0  0  1 |
```

che risulta invertibile (triangolare inferiore con ingressi diagonali non nulli).

Questo conclude la dimostrazione del teorema. □

Osservazione. La dimostrazione appena vista si appoggia pesantemente sul fatto che gli spazi in questione hanno dimensione 2. Già nel caso del tensore $\Psi_{\langle 2,2,3\rangle}$ usando lo stesso argomento proposto i flattening di Koszul non danno informazioni migliori dei flattening standard (in particolare per ottenere una stima significativa del rango le dimensioni degli spazi U, V, W sono troppo piccole). Tuttavia, come esposto in [13] è comunque possibile, utilizzando opportuni flattening di Koszul, ottenere informazioni sul rango bordo per dimensioni superiori.

Macaulay2. Vediamo nel seguente script una funzione per calcolare i flattening di Koszul utilizzati nel Teorema 3.4.2, insieme ai passaggi utilizzati nella dimostrazione.

```
kFlat = (t, U, V, W) -> (
    xr := res(U);
    yy := gens(V);
    zz := gens(W);
    return diff(yy, diff(transpose(zz), diff(xr.dd_2, t)));
    )

n = 2
R = QQ[x_(0, 0) .. x_(n - 1, n - 1),
       y_(0, 0) .. y_(n - 1, n - 1),
       z_(0, 0) .. z_(n - 1, n - 1),
       a .. d]
S = QQ[a .. d]
use R
A = ideal(x_(0, 0) .. x_(n - 1, n - 1))
B = ideal(y_(0, 0) .. y_(n - 1, n - 1))
C = ideal(z_(0, 0) .. z_(n - 1, n - 1))
t = sum(n, i -> sum(n, j -> sum(n, k -> x_(i, j) * y_(j, k) * z_(k, i))))

t1 = t -
     sum(n, i -> x_(i, i)) *
     sum(n, i -> y_(i, i)) *
     (a * z_(0, 0) + b * z_(0, 1) + c * z_(1, 0) + d * z_(1, 1))
rr1 = kFlat(t1, A, B, C)
rr11 = submatrix(rr1, , {11, 1, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 4, 12})
rr12 = submatrix(rr1, , {11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 4, 12})
rr13 = submatrix(rr1, , {11, 1, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 0, 12})
rr14 = submatrix(rr1, , {11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 15, 16, 14, 0, 12})

t2 = t -
     (a * x_(0, 0) + b * x_(0, 1) + c * x_(1, 0) + d * x_(1, 1)) *
     y_(0, 0) *
     z_(0, 0)
rr2 = kFlat(t2, A, B, C)
rr21 = submatrix(rr2, , {11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 1, 16, 14, 4, 12})
rr22 = submatrix(rr2, , {11, 23, 8, 7, 9, 20, 6, 18, 5, 17, 2, 1, 16, 14, 0, 12})

t3 = t -
```

```
     (a * x_(0, 0) + b * x_(0, 1) + c * x_(1, 0) + d * x_(1, 1)) *
     y_(0, 0) *
     z_(1, 0)
rr3 = kFlat(t3, A, B, C)
rr31 = submatrix(rr3, , {12, 0, 1, 3, 15, 2, 17, 5, 18, 6, 7, 9, 21, 8, 23, 11})
```

Bibliografia

[1] P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic complexity theory.* Grundlehren der mathematischen Wissenschaften, No. 315 Springer Verlag 1997.

[2] C. Ciliberto. *Algebra Lineare.* Bollati Boringhieri, 1994.

[3] H. F. de Groote. *On varieties of optimal algorithms for the computation of bilinear mappings I. Optimal algorithms for 2 $\times$ 2-matrix multiplication.* Theoretical Computer Science, Volume 7, Issue 1, 1978, Pages 1-24.

[4] H. F. de Groote. *On varieties of optimal algorithms for the computation of bilinear mappings II. Optimal algorithms for 2 $\times$ 2-matrix multiplication.* Theoretical Computer Science, Volume 7, Issue 2, 1978, Pages 127-148

[5] J. Harris. *Algebraic Geometry.* Springer, 1992.

[6] J. Hastad. *Tensor Rank is NP-complete.* Journal of Algorithms, Volume 11, Issue 4, December 1990, Pages 644–654.

[7] J. D. Hauenstein, C. Ikenmeyer, and J.M. Landsberg *Equations for lower bounds on border rank.* Exp. Math., 22:4 (2013), 372–383.

[8] J. E. Hopcroft and L. R. Kerr. *On minimizing the number of multiplications necessary for matrix multiplication*, SIAM J. Appl. Math. 20 (1971), 30–36.

[9] C. Hillar and L. Lim. *Most tensor problems are NP-hard.* arXiv:0911.1393

[10] J.M. Landsberg, *Tensors: Geometry and Applications.* American Mathematical Society, 2012.

[11] J.M. Landsberg, *The border rank of the multiplication of* 2×2 *matrices is seven.* J. Amer. Math. Soc. 19 (2006), 447-459.

[12] J.M. Landsberg and G. Ottaviani, *New lower bounds for the border rank of matrix multiplication.* arXiv:1112.6007 to appear in Theory of Computing.

[13] G. Ottaviani, *Complexity of Matrix Multiplication and Tensor Rank.* KIAS, Seoul, February 27, 2014. `http://web.math.unifi.it/users/ottaviani/colloquium_kias.pdf`.

[14] V. Strassen. *Gaussian Elimination is not Optimal.* Numer. Math. 13 1969 354–356.

[15] S. Winograd, *On multiplication of* 2×2 *matrices*, Linear Algebra and Appl. 4 (1971), 381–388.