



UNIVERSITÀ DEGLI STUDI DI FERRARA

**FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E
NATURALI**

Corso di Laurea Triennale in Matematica
Indirizzo Modelli Matematici per l'Economia

**TEORIA SPETTRALE DI
SERIE TEMPORALI STAZIONARIE**

Relatore:
**Chiar.mo Prof.
Josef Eschgfäller**

Laureanda:
Eleonora Loss

Anno Accademico 2009-2010

Indice

1. Preliminari e notazioni	3
2. Serie temporali stazionaria	8
3. La misura spettrale	18
4. Spazi vettoriali con un prodotto scalare	31
5. Approssimazione in uno spazio di Banach	38
6. Operatori lineari continui	44
7. Spazi di Hilbert	45
8. Il teorema di Fréchet-Riesz	52
9. Il rapporto di Rayleigh	56
10. Sistemi ortonormali completi	58
11. Ortogonalizzazione	64
12. Predizione lineare	66
13. Matrici di Toeplitz	68
Bibliografia	74

1. Preliminari e notazioni

Situazione 1.1. In tutta la tesi (Ω, \mathcal{A}, p) sia uno spazio di probabilità finito e in genere sottinteso.

Per le dimostrazioni degli enunciati in questo capitolo rimandiamo ai corsi di analisi e di calcolo delle probabilità.

Per applicazioni $X : \Omega \rightarrow \Omega'$ useremo, per $B \subset \Omega'$, la notazione

$$(X \in B) := \{\omega \in \Omega \mid X(\omega) \in B\} = X^{-1}(B)$$

e, nel caso $\Omega' = \mathbb{R}$,

$$(X \leq a) := \{\omega \in \Omega \mid X(\omega) \leq a\}$$

come d'uso nel calcolo delle probabilità.

Definizione 1.2. Un'applicazione $X : \Omega \rightarrow \mathbb{R}$ si dice *misurabile*, se è \mathcal{A} -Borel(\mathbb{R})-misurabile; similmente un'applicazione $X : \Omega \rightarrow \mathbb{C}$ si dice *misurabile*, se è \mathcal{A} -Borel(\mathbb{C})-misurabile.

Identificheremo in maniera naturale le applicazioni $X : \Omega \rightarrow \mathbb{R}$ con le applicazioni $X : \Omega \rightarrow \mathbb{C}$ con $X(\omega) \in \mathbb{R}$; per esse i due concetti di misurabilità evidentemente coincidono.

Definizione 1.3. Una *variabile aleatoria reale* è un'applicazione misurabile $X : \Omega \rightarrow \mathbb{R}$, una *variabile aleatoria complessa* è un'applicazione misurabile $X : \Omega \rightarrow \mathbb{C}$.

Nel senso della def. 1.2 le variabili aleatorie reali sono identificate con le variabili aleatorie complesse a valori reali.

Definizione 1.4. Per l'integrale (quando esiste) di una funzione f rispetto a una misura μ su un insieme A utilizziamo la notazione

$$\int_A f(a) d\mu(a)$$

Definizione 1.5. Per $k > 0$ siano

$$\mathcal{L}_{\mathbb{R}}^k := \mathcal{L}_{\mathbb{R}}^k(\Omega, \mathcal{A}, p) := \left\{ X : \Omega \rightarrow \mathbb{R} \mid X \text{ misurabile e } \int_{\Omega} |X(\omega)|^k dp(\omega) < \infty \right\}$$

$$\mathcal{L}^k := \mathcal{L}^k(\Omega, \mathcal{A}, p) := \left\{ X : \Omega \rightarrow \mathbb{C} \mid X \text{ misurabile e } \int_{\Omega} |X(\omega)|^k dp(\omega) < \infty \right\}$$

In genere ci limiteremo ai casi $k = 1, 2$.

Si osservi che $X \in \mathcal{L}^k \iff X^k \in \mathcal{L}^1$. Utilizzando la disuguaglianza di Hölder si dimostra (nella nostra ipotesi che p sia una misura di probabilità) che $\mathcal{L}^2 \subset \mathcal{L}^1$.

Definizione 1.6. (Ω', \mathcal{B}) sia uno spazio misurabile ed $X : \Omega \rightarrow \Omega'$ un'applicazione \mathcal{A} - \mathcal{B} -misurabile. Per $B \in \mathcal{B}$ poniamo allora

$$p[X](B) := p(X \in B)$$

Si dimostra facilmente che $p[X]$ è una misura di probabilità su \mathcal{B} .

Teorema 1.7 (formula astratta di trasformazione). (Ω', \mathcal{B}) sia uno spazio misurabile e $X : \Omega \rightarrow \Omega'$ un'applicazione \mathcal{A} - \mathcal{B} -misurabile. Sia data una funzione misurabile $g : \Omega' \rightarrow \mathbb{C}$ e sia $k > 0$. Allora:

$$(1) g \circ X \in \mathcal{L}^k(\Omega, \mathcal{A}, p) \iff g \in \mathcal{L}^k(\Omega', \mathcal{B}, p[X]).$$

$$(2) g \circ X \in \mathcal{L}^k(\Omega, \mathcal{A}, p) \implies \int_{\Omega} g(X(\omega)) dp(\omega) = \int_{\Omega'} g(\omega') dp[X](\omega').$$

Definizione 1.8. Sia $X \in \mathcal{L}^1$. Allora

$$MX := \int_{\Omega} X(\omega) dp(\omega)$$

si chiama la *media* di X .

Osservazione 1.9. Siano $X, Y \in \mathcal{L}^2$. Allora $X + Y, XY, X\bar{Y} \in \mathcal{L}^1$.

In particolare si ha $|X|^2 \in \mathcal{L}^1$.

È chiaro che ciò implica anche che $(X - MX) \overline{(Y - MY)} \in \mathcal{L}^1$ e $|X - MX|^2 \in \mathcal{L}^1$.

Definizione 1.10. Siano $X, Y \in \mathcal{L}^2$. Allora

$$\text{cov}(X, Y) := M(X - MX) \overline{(Y - MY)}$$

si chiama la *covarianza* di X ed Y .

Si osservi che $\text{cov}(X, Y) = \overline{\text{cov}(Y, X)}$.

Definizione 1.11. Sia $X \in \mathcal{L}^2$.

$$\sigma^2(X) := \text{cov}(X, X) = M|X - MX|^2$$

si chiama la *varianza* di X .

La *deviazione standard* $\sigma(X)$ è definita come la radice quadrata (≥ 0) della varianza.

Lemma 1.12. Siano $X, Y \in \mathcal{L}^2$

$$\text{Allora } \text{cov}(X, Y) = MX\bar{Y} - MX \cdot \overline{MY}.$$

Dimostrazione. Infatti

$$\begin{aligned} \text{cov}(X, Y) &= M(X - MX) \overline{(Y - MY)} = M(X - MX) (\bar{Y} - \overline{MY}) \\ &= M(X\bar{Y} - X \cdot \overline{MY} - MX \cdot \bar{Y} + MX \cdot \overline{MY}) \\ &= MX\bar{Y} - MX \cdot \overline{MY} - MX \cdot \bar{Y} + MX \cdot \overline{MY} \\ &= MX\bar{Y} - MX \cdot \overline{MY} \end{aligned}$$

perché naturalmente $M\bar{Y} = \overline{MY}$.

Corollario 1.13. Sia $X \in \mathcal{L}^2$. Allora $\sigma^2(X) = M|X|^2 - |MX|^2$.

Proposizione 1.14.

(1) Siano $X_1, \dots, X_r, Y_1, \dots, Y_r \in \mathcal{L}^2$ ed $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_r \in \mathbb{C}$. Allora

$$\text{cov} \left(\sum_{j=1}^r \alpha_j X_j, \sum_{k=1}^r \beta_k Y_k \right) = \sum_{j=1}^r \sum_{k=1}^r \alpha_j \overline{\beta_k} \text{cov} (X_j, Y_k)$$

(2) Siano $X, Y \in \mathcal{L}^2$. Allora

$$\sigma^2 (X + Y) = \sigma^2 (X) + \sigma^2 (Y) + 2 \text{Re} \text{cov} (X, Y).$$

Dimostrazione. (1) Tralasciando i limiti di sommazione abbiamo

$$\begin{aligned} & \text{cov} (\alpha_1 X_1 + \dots + \alpha_r X_r, \beta_1 Y_1 + \dots + \beta_r Y_r) \\ &= M \sum_j \sum_k \alpha_j X_j \overline{\beta_k Y_k} - \sum_j \alpha_j M X_j \sum_k \overline{\beta_k M Y_k} \\ &= \sum_j \sum_k \alpha_j \overline{\beta_k} (M X_j Y_k - M X_j \overline{M Y_k}) \\ &= \sum_j \sum_k \alpha_j \overline{\beta_k} \text{cov} (X_j, Y_k) \end{aligned}$$

(2)

$$\begin{aligned} \sigma^2 (X + Y) &= \text{cov} (X + Y, X + Y) \\ &= \text{cov} (X, X) + \text{cov} (X, Y) + \text{cov} (Y, X) + \text{cov} (Y, Y) \\ &= \sigma^2 (X) + \text{cov} (X, Y) + \text{cov} (Y, X) + \sigma^2 (Y) \end{aligned}$$

Ma $\text{cov} (X, Y) + \text{cov} (Y, X) = \text{cov} (X, Y) + \overline{\text{cov} (X, Y)} = 2 \text{Re} \text{cov} (X, Y)$.

Definizione 1.15. Le variabili aleatorie complesse X ed Y si dicono *indipendenti*, se per ogni $A, B \in \text{Borel} (\mathbb{C})$ vale

$$p (X \in A, Y \in B) = p (X \in A) \cdot p (Y \in B)$$

Proposizione 1.16. Le variabili aleatorie complesse X ed Y siano *indipendenti*.

Allora $\text{cov} (X, Y) = 0$ e quindi $\sigma^2 (X + Y) = \sigma^2 (X) + \sigma^2 (Y)$.

Definizione 1.17. Due variabili aleatorie complesse X ed Y si dicono *identicamente distribuite* se per ogni $A \in \text{Borel} (\mathbb{C})$ vale

$$p (X \in A) = p (Y \in A).$$

La stessa definizione si applica a più di due variabili aleatorie.

Proposizione 1.18. X ed Y siano variabili aleatorie complesse. Allora sono equivalenti:

- (1) X ed Y sono *identicamente distribuite*.
- (2) $p (\text{Re} X \leq a, \text{Im} X \leq b) = p (\text{Re} Y \leq a, \text{Im} Y \leq b)$ per ogni $a, b \in \mathbb{R}$.

Dimostrazione. L'enunciato discende dal fatto che

$$\text{Borel} (\mathbb{C}) = \sigma\text{-algebra} (\{(-\infty, a] \times (-\infty, b] \mid a, b \in \mathbb{R}\})$$

Proposizione 1.19. X ed Y siano variabili aleatorie complesse *identicamente distribuite*. Allora:

- (1) Se $X \in \mathcal{L}^1$, allora anche $Y \in \mathcal{L}^1$ e si ha $MX = MY$.
- (2) Se $X \in \mathcal{L}^2$, allora anche $Y \in \mathcal{L}^2$ e si ha $\sigma^2(X) = \sigma^2(Y)$.

Dimostrazione. Ciò segue dalla definizione dell'integrale.

Definizione 1.20. La distribuzione di probabilità di una variabile aleatoria reale X è la funzione $\bigcirc_t p(X \leq t) : \mathbb{R} \rightarrow [0, 1]$.

Proposizione 1.21. X ed Y siano due variabili aleatorie reali. Allora sono equivalenti:

- (1) X ed Y sono identicamente distribuite.
- (2) X ed Y possiedono la stessa distribuzione di probabilità.

Dimostrazione. Si ottiene come caso speciale della prop. 1.18 oppure dal fatto che $Borel(\mathbb{R}) = \sigma\text{-algebra}(\{(-\infty, a] \mid a \in \mathbb{R}\})$.

Proposizione 1.22. X ed Y siano variabili aleatorie reali. Allora sono equivalenti:

- (1) X ed Y sono indipendenti.
- (2) $p(X \leq a, Y \leq b) = p(X \leq a)p(Y \leq b)$ per ogni $a, b \in \mathbb{R}$.

Dimostrazione. Ciò segue di nuovo da $Borel(\mathbb{R}) = \sigma\text{-algebra}(\{(-\infty, a] \mid a \in \mathbb{R}\})$.

Osservazione 1.23. X sia una variabile aleatoria reale. Allora:

- (1) Se $X \in \mathcal{L}^1$, allora $MX = \int_{\mathbb{R}} t dp[X](t)$.
- (2) Se $X \in \mathcal{L}^1$ e $\mu := MX$, allora $\sigma^2(X) = \int_{\mathbb{R}} (t - \mu)^2 dp[X](t)$.

Dimostrazione. Ciò segue direttamente dal teorema 1.7.

Definizione 1.24. Una serie temporale (unodimensionale) è una successione $\bigcirc_{n \in \mathbb{Z}} X_n$ di variabili aleatorie complesse.

Una serie temporale $\bigcirc_n X_n$ si dice *reale*, se X_n è reale per ogni n . Con un leggero abuso di notazione scriveremo talvolta $\bigcirc_n X_n \in \mathcal{L}^k$ se $X_n \in \mathcal{L}^k$ per ogni n .

Definizione 1.25. Un rumore bianco è una serie temporale $\bigcirc_n B_n$ con le seguenti proprietà :

- (1) $B_n \in \mathcal{L}^2$ per ogni n .
- (2) $MB_n = 0$ per ogni n .
- (3) La varianza $\sigma^2(B_n) =: \sigma^2$ non dipende da n .
- (4) $\sigma^2 > 0$.
- (5) $\text{cov}(B_n, B_m) = 0$ per $n \neq m$.

Diremo allora che $\bigcirc_n B_n$ è un RB (σ^2).

Osservazione 1.26. $\bigcirc_n B_n$ sia una serie temporale per la quale siano soddisfatte le seguenti condizioni:

- (1) $B_n \in \mathcal{L}^2$ per ogni n .
- (2) Le variabili aleatorie B_n sono identicamente distribuite.
- (3) La loro media comune è uguale a zero.
- (4) La loro varianza comune è maggiore di zero.
- (5) Le variabili aleatorie B_n sono indipendenti a due a due.

Allora $\bigcirc_n B_n$ è un RB (σ^2).

Dimostrazione. Affinché la serie temporale $\bigcirc_n B_n$ sia un RB (σ^2) devono valere le condizioni della def. 1.25:

- (1) Per ipotesi.
- (2) Per ipotesi. Che la media non dipende da n segue dalla prop. 1.19.
- (3) La prop. 1.19 implica anche che la varianza non dipende da n .
- (4) Per ipotesi.
- (5) Dalla condizione (5) e per la prop. 1.16 segue che $\text{cov}(B_n, B_m) = 0$ per $n \neq m$.

Definizione 1.27. Su \mathcal{L}^2 possiamo introdurre un *semiprodotto scalare* con

$$\|X, Y\| := \int_{\Omega} X(\omega) \overline{Y(\omega)} dp(\omega) = MX\overline{Y}$$

che possiede tutte le proprietà di un prodotto scalare (complesso) tranne che $\|X, X\| = 0$ non implica $X = 0$.

Identificando due variabili aleatorie X ed Y se $p(X = Y) = 1$, otteniamo uno spazio di Hilbert L^2 . Quando necessario, la classe di equivalenza di X in L^2 viene denotata con $(X, \text{in } L^2)$. Cfr. Weidmann, pag. 25.

Osservazione 1.28. Con la notazione della def. 1.27 per $X, Y \in \mathcal{L}^2$ abbiamo

$$\begin{aligned} MX &= \|X, \bigcirc_{\omega} 1\| \\ \text{cov}(X, Y) &= \|X - MX, Y - MY\| \\ \|X\|^2 &= \sigma^2(X) + |MX|^2 \end{aligned}$$

Qui la seminorma $\|X\| := \sqrt{\|X, X\|} = \sqrt{M|X|^2}$ è definita rispetto al semiprodotto scalare, per cui la terza equazione segue dal cor. 1.13.

2. Serie temporali stazionarie

Situazione 2.1. Sia data una serie temporale $\bigcirc_n X_n \in \mathcal{L}^2$.

Definizione 2.2. La serie temporale $\bigcirc_n X_n$ si dice

- (1) *strettamente stazionaria*, se per ogni $k \in \mathbb{N} + 1$ e per ogni m ed ogni $t_1, \dots, t_k \in \mathbb{Z}$ le variabili aleatorie k -dimensionali $(X_{t_1}, \dots, X_{t_k})$ e $(X_{t_1+m}, \dots, X_{t_k+m})$ sono identicamente distribuite, cioè se $p(X_{t_1+m} \leq a_1, \dots, X_{t_k+m} \leq a_k) = p(X_{t_1} \leq a_1, \dots, X_{t_k} \leq a_k)$ per ogni $a_1, \dots, a_k \in \mathbb{R}$;
- (2) *stazionaria*, se $\mu := MX_t$ non dipende da t e $\text{cov}(X_{t+m}, X_t)$ per ogni $m \in \mathbb{Z}$ fissato non dipende da t .
In particolare quindi $\sigma_X^2 := \text{cov}(X_t, X_t) = \sigma^2(X_t)$ non dipende da t .

Osservazione 2.3. Se $\bigcirc_n X_n$ è strettamente stazionaria, allora è anche stazionaria.

Osservazione 2.4. Le seguenti condizioni sono equivalenti:

- (1) Per ogni $m \in \mathbb{Z}$ fissato $\text{cov}(X_{t+m}, X_t)$ non dipende da t .
- (2) Per ogni $m, t, s \in \mathbb{Z}$ vale $\text{cov}(X_{t+m}, X_{s+m}) = \text{cov}(X_t, X_s)$.

Dimostrazione. (1) \implies (2):

$$\begin{aligned} \text{cov}(X_{t+m}, X_{s+m}) &= \text{cov}(X_{s+(t-s)+m}, X_{s+m}) \\ &= \text{cov}(X_{s+(t-s)}, X_s) = \text{cov}(X_t, X_s) \end{aligned}$$

(2) \implies (1) : $\text{cov}(X_{t+m}, X_t) = \text{cov}(X_m, X_0)$ non dipende da t .

Definizione 2.5. La serie temporale $\bigcirc_n X_n$ sia stazionaria. Allora possiamo definire una funzione $\gamma : \mathbb{Z} \longrightarrow \mathbb{C}$ ponendo

$$\gamma(m) := \text{cov}(X_m, X_0)$$

Per ipotesi allora $\gamma(m) = \text{cov}(X_{t+m}, X_t)$ per ogni $m, t \in \mathbb{Z}$.

γ si chiama la *funzione di autocovarianza* di $\bigcirc_n X_n$.

Sia $\mu := MX_0$ (uguale a MX_t per ogni t). Allora diciamo che $\bigcirc_n X_n$ è una STS (μ, γ) .

Osservazione 2.6. $\bigcirc_n X_n$ sia una STS (μ, γ) . Usando la notazione della def. 1.27 per ogni $m, t \in \mathbb{Z}$ abbiamo

$$\gamma(m) = \|X_m - \mu, X_0 - \mu\| = \|X_{t+m} - \mu, X_t - \mu\|$$

Osservazione 2.7. $\bigcirc_n X_n$ sia una STS (μ, γ) . Per ogni $t, s \in \mathbb{Z}$ allora

$$\text{cov}(X_t, X_s) = \gamma(t - s)$$

Dimostrazione. Per l'oss. 2.4 abbiamo

$$\text{cov}(X_t, X_s) = \text{cov}(X_{t-s}, X_0) = \gamma(t-s)$$

Osservazione 2.8. $\bigcirc_n X_n$ sia una STS (μ, γ) . Allora $\gamma(0) = \sigma^2(X_t)$ per ogni $t \in \mathbb{Z}$. In particolare $\gamma(0)$ è reale e ≥ 0 .

Definizione 2.9. $\bigcirc_n X_n$ sia una STS (μ, γ) con $\gamma(0) > 0$.

Allora la funzione $\bigcirc_m \gamma(m)/\gamma(0) : \mathbb{Z} \rightarrow \mathbb{C}$ si chiama la *funzione di autocorrelazione* di $\bigcirc_n X_n$.

Osservazione 2.10. $\bigcirc_n X_n$ sia una STS (μ, γ) . Allora per ogni $t_1, \dots, t_k, m \in \mathbb{Z}$ si ha

$$\gamma(m) = M \frac{X_{t_1+m} \bar{X}_{t_1} + \dots + X_{t_k+m} \bar{X}_{t_k}}{k}$$

Proposizione 2.11. $\bigcirc_n X_n$ sia una STS (μ, γ) .

Allora $\gamma(-m) = \overline{\gamma(m)}$ per ogni $m \in \mathbb{Z}$.

Se la serie temporale $\bigcirc_n X_n$ è reale, abbiamo quindi $\gamma(-m) = \gamma(m)$ per ogni $m \in \mathbb{Z}$.

Dimostrazione. Usando l'oss. 2.6 abbiamo

$$\begin{aligned} \gamma(-m) &= \|X_{-m} - \mu, X_0 - \mu\| = \|X_0 - \mu, X_m - \mu\| \\ &= \overline{\|X_m - \mu, X_0 - \mu\|} = \overline{\gamma(m)} \end{aligned}$$

Esempio 2.12. $\bigcirc_n X_n$ sia un RB (σ^2) . Allora $\bigcirc_n X_n$ è una STS $(0, \gamma)$ con

$$\gamma(m) = \begin{cases} 0 & \text{per } m \neq 0 \\ \sigma^2 & \text{per } m = 0 \end{cases}$$

La funzione di autocorrelazione è quindi data da

$$\rho(m) = \begin{cases} 0 & \text{per } m \neq 0 \\ 1 & \text{per } m = 0 \end{cases}$$

Dimostrazione. Per ipotesi $\bigcirc_n X_n$ è un RB (σ^2) , quindi $MX_n = 0$ per ogni n . Per $m \neq 0$ inoltre $\text{cov}(X_{t+m}, X_t) = 0$ per ogni t .

Infine $\text{cov}(X_t, X_t) = \sigma^2$.

Definizione 2.13. Sia $q \in \mathbb{N} + 1$. $\bigcirc_n X_n$ si dice una serie temporale a media mobile di ordine q , se esistono un rumore bianco $B = \bigcirc_n B_n$ e coefficienti costanti $\theta_1, \dots, \theta_q \in \mathbb{C}$ con $\theta_q \neq 0$ tali che

$$X_t = B_t + \theta_1 B_{t-1} + \dots + \theta_q B_{t-q}$$

per $t \in \mathbb{Z}$. Diciamo allora che $\bigcirc_n B_n$ è una serie temporale di tipo MM $(\theta_1, \dots, \theta_q, B)$. Come nella def. 2.2 denotiamo con σ_B^2 la varianza del rumore bianco.

Nella terminologia inglese una tale serie temporale si dice di tipo *moving average (MA)*.

Proposizione 2.14. $\bigcirc_n X_n$ sia di tipo MM $(\theta_1, \dots, \theta_q, B)$. Allora $\bigcirc_n X_n$ è una STS $(0, \gamma)$ con

$$\gamma(m) = \begin{cases} \sigma_B^2 \cdot (\theta_m + \theta_{m+1}\bar{\theta}_1 + \dots + \theta_q\bar{\theta}_{q-m}) & \text{per } 0 \leq m \leq q \\ 0 & \text{per } m > q \\ \overline{\gamma(-m)} & \text{per } m < 0 \end{cases}$$

dove abbiamo posto $\theta_0 := 1$.

Dimostrazione. (1) È immediato che $X_n \in \mathcal{L}^2$ e $MX_n = 0$ per ogni n .

(2) Ponendo $\theta_j := 0$ per $j \notin \{0, 1, \dots, q\}$ e sommando da $-\infty$ a ∞ , abbiamo

$$\begin{aligned} \text{cov}(X_{t+m}, X_t) &= \text{cov}\left(\sum_j \theta_j B_{t+m-j}, \sum_k \theta_k B_{t-k}\right) \\ &= \sigma_B^2 \sum_j \sum_k \theta_j \bar{\theta}_k \delta_{m-j, -k} = \sigma_B^2 \sum_j \theta_j \bar{\theta}_{j-m} \\ &= \sigma_B^2 \sum_j \theta_{j+m} \bar{\theta}_j \end{aligned}$$

Questa somma non dipende da t e quindi la serie temporale $\bigcirc_n X_n$ è stazionaria.

(3) Per $m \geq 0$ otteniamo $\sigma_B^2 (\theta_m \bar{\theta}_0 + \theta_{m+1} \bar{\theta}_1 + \dots + \theta_{m+q} \bar{\theta}_q)$

Questa somma si annulla per $m > q$.

Per $m < 0$ usiamo la prop. 2.11.

Corollario 2.15. $\bigcirc_n X_n$ sia di tipo MM $(\theta_1, \dots, \theta_q, B)$. Allora la varianza comune delle variabili aleatorie X_n è data da

$$\sigma_X^2 = \gamma(0) = (1 + |\theta_1|^2 + \dots + |\theta_q|^2) \sigma_B^2$$

Corollario 2.16. $\bigcirc_n X_n$ sia di tipo MM (θ, B) con $\theta \in \mathbb{R}$. Allora

$$\gamma(m) = \begin{cases} (1 + \theta^2) \sigma_B^2 & \text{per } m = 0 \\ \theta \cdot \sigma_B^2 & \text{per } m = \pm 1 \\ 0 & \text{altrimenti} \end{cases}$$

Nota 2.17. $\bigcirc_n X_n$ sia di tipo MM (θ, B) con $\theta \in \mathbb{R}$. Assumiamo di conoscere (in genere approssimativamente mediante considerazioni em-

piriche) $\gamma(0) =: \gamma_0$ e $\gamma(1) =: \gamma_1$. Vogliamo calcolare i parametri θ e σ_B^2 , entrambi $\neq 0$ per definizione.

Poniamo $\rho_1 := \rho(1) = \frac{\gamma_1}{\gamma_0} \neq 0$.

Dal cor. 2.16 abbiamo

$$\begin{aligned}\sigma_B^2(1 + \theta^2) &= \gamma_0 \\ \sigma_B^2\theta &= \gamma_1\end{aligned}$$

Ciò implica $\frac{\theta}{1 + \theta^2} = \frac{\gamma_1}{\gamma_0} = \rho_1$, per cui

$$\rho_1\theta^2 - \theta + \rho_1$$

cosicché

$$\theta = \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1}$$

Inoltre $\sigma_B^2 = \frac{\gamma_1}{\theta}$ e siccome $\frac{1 + \theta^2}{\theta} = \frac{1}{\rho_1}$ implica $\frac{1}{\theta} = \frac{1}{\rho_1} - \theta$,

otteniamo

$$\sigma_B^2 = \gamma_1 \left(\frac{1}{\rho_1} - \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1} \right) = \frac{\gamma_1}{\rho_1} \left(\frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2} \right)$$

e quindi

$$\sigma_B^2 = \gamma_0 \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2}$$

Possiamo perciò distinguere i seguenti casi:

(1) Se $|\rho_1| > \frac{1}{2}$, il problema non ha soluzione (perché abbiamo postulato che $\theta \in \mathbb{R}$).

(2) Per $\rho_1 = \frac{1}{2}$ otteniamo $\theta = 1$ e $\sigma_B^2 = \frac{\gamma_0}{2}$.

(3) Per $\rho_1 = -\frac{1}{2}$ otteniamo $\theta = -1$ e $\sigma_B^2 = \frac{\gamma_0}{2}$.

(4) Se $|\rho_1| < \frac{1}{2}$ otteniamo due soluzioni distinte.

Cfr. Neusser, pag 18-19.

Definizione 2.18. La serie temporale $\bigcirc_n X_n$ si dice *autoregressiva di ordine 1*, se sono soddisfatte le seguenti condizioni:

(1) $\bigcirc_n X_n$ è stazionaria.

(2) Esistono $a \in \mathbb{C} \setminus 0$ e un rumore bianco B tali che

$$X_t - aX_{t-1} = B_t$$

per ogni $t \in \mathbb{Z}$

(3) $\text{cov}(X_t, B_s) = 0$ per $t, s \in \mathbb{Z}$ con $s > t$ (indipendenza dal futuro).

Diremo allora che $\bigcirc_n X_n$ è di tipo AR(a, B).

Osservazione 2.19. $\bigcirc_n X_n$ sia di tipo AR(a, B).

Allora $0 < |a| < 1$ e $\gamma(0) = \frac{\sigma_B^2}{1 - |a|^2}$.

Dimostrazione. Per ipotesi $a \neq 0$.

Abbiamo in particolare $X_1 = aX_0 + B_1$ e inoltre $\text{cov}(X_0, B_1) = 0$ per la condizione (3) nella def. 2.18.

Dalla sesquilinearità della covarianza (cfr. oss. 1.28) otteniamo

$$\begin{aligned} \gamma(0) &= \text{cov}(X_1, X_1) = \text{cov}(aX_0 + B_1, aX_0 + B_1) \\ &= a\bar{a} \text{cov}(X_0, X_0) + \text{cov}(B_1, B_1) \\ &= |a|^2 \gamma(0) + \sigma_B^2 \end{aligned}$$

e quindi $(1 - |a|^2) \gamma(0) = \sigma_B^2$.

Siccome $\sigma_X^2 > 0$, ciò implica $1 - |a|^2 \neq 0$, per cui $\gamma(0) = \frac{\sigma_B^2}{1 - |a|^2}$.

Però $\gamma(0)$ deve essere reale e ≥ 0 e ciò implica $|a| < 1$.

Osservazione 2.20. $\bigcirc_n X_n$ sia di tipo AR(a, B).

Allora $MX_n = 0$ per ogni n .

Dimostrazione. Per la stazionarietà $\mu := MX_n$ non dipende da n .

Per ipotesi $X_1 = aX_0 + B_1$ con $MB_1 = 0$, per cui

$$\mu = MX_1 = aMX_0 = a\mu$$

Però $a \neq 1$ per l'oss. 2.19 e ciò implica $\mu = 0$.

Osservazione 2.21. Siano B un rumore bianco e $a \in \mathbb{C}$ con $0 < |a| < 1$. Non sappiamo ancora se esiste una serie temporale di tipo AR(a, B). Assumiamo però che $\bigcirc_n X_n$ sia di tipo AR(a, B).

Allora per ogni $t \in \mathbb{Z}$ per ogni $n \in \mathbb{N}$ si ha

$$X_t = a^n X_{t-n} + \sum_{k=0}^{n-1} a^k B_{t-k} \quad (*)$$

Perciò

$$\|X_t - \sum_{k=0}^{n-1} a^k B_{t-k}\| = |a|^n \|X_{t-n}\| = |a|^n \sigma_X$$

perché $MX_{t-n} = 0$ per l'oss. 2.20.

Siccome $|a| < 1$ per l'oss. 2.19, ciò implica che $X_t \stackrel{L^2}{=} \sum_{k=0}^{\infty} a^k B_{t-k}$.

Nel prossimo teorema dimostreremo che in questo modo otteniamo veramente una serie temporale di tipo AR (a, B) .

Dimostrazione. Dobbiamo solo dimostrare la relazione (*). Essa si ottiene applicando ripetutamente la ricorrenza al punto (2) della def. 2.18:

$$\begin{aligned} X_t &= aX_{t-1} + B_t \\ &= a(aX_{t-2} + B_{t-1}) + B_t = a^2X_{t-2} + aB_{t-1} + B_t \\ &= a^2(aX_{t-3} + B_{t-2}) + aB_{t-1} + B_t = a^3X_{t-3} + a^2B_{t-2} + aB_{t-1} + B_t \end{aligned}$$

ecc.

Teorema 2.22. Siano B un rumore bianco ed $a \in \mathbb{C}$ con $0 < |a| < 1$.

Allora la serie temporale $\bigcirc_t X_t$ definita da $X_t := \sum_{k=0}^{\infty} a^k B_{t-k}$ è ben definita ed è (in L^2) l'unica serie temporale di tipo AR (a, B) .

$$\text{Per } m \in \mathbb{N} \text{ si ha } \gamma(m) = \sigma_B^2 \frac{a^m}{1 - |a|^2}.$$

Dimostrazione. In questa dimostrazione lavoriamo in L^2 , senza distinguere esplicitamente tra una funzione $X \in \mathcal{L}^2$ e la sua classe $(X, \text{ in } L^2)$.

(1) L'unicità deriva dall'oss. 2.21.

(2) È chiaro che per t fissato la serie $\bigcirc_t \sum_{k=0}^n a^k B_{t-k}$ è sempre una serie di Cauchy in L^2 , per cui il limite X_t è un ben definito elemento di L^2 .

(3) Dimostriamo la relazione di ricorrenza per ogni t :

$$X_t - aX_{t-1} = \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n a^k B_{t-k} - a \sum_{k=0}^{n-1} a^k B_{t-k-1} \right) = B_t$$

(4) Siano $t, s \in \mathbb{Z}$ con $s > t$. Allora per la continuità del prodotto scalare

$$\text{cov}(X_t, B_s) = \text{cov} \left(\sum_{k=0}^{\infty} a^k B_{t-k}, B_s \right) = \sum_{k=0}^{\infty} a^k \text{cov}(B_{t-k}, B_s) = 0$$

perché $s > t - k$ per ogni $k \in \mathbb{N}$.

(5) Rimane da dimostrare la stazionarietà della serie temporale $\bigcirc_t X_t$.

Utilizzando ad esempio la prima equazione nell'oss. 1.28 è chiaro che $MX_t = 0$ per ogni t . Siano $t \in \mathbb{Z}$ e $m \in \mathbb{N}$. Allora

$$\begin{aligned}
\text{cov}(X_{t+m}, X_t) &= \text{cov}\left(\sum_{k=0}^{\infty} a^k B_{t+m-k}, \sum_{j=0}^{\infty} a^j B_{t-j}\right) \\
&= \sigma_B^2 \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \delta_{t+m-k, t-j} a^k \bar{a}^j \\
&= \sigma_B^2 \sum_{j=0}^{\infty} a^{j+m} \bar{a}^j = \sigma_B^2 a^m \sum_{j=0}^{\infty} |a|^{2j} = \sigma_B^2 \frac{a^m}{1 - |a|^2}
\end{aligned}$$

Quindi $\text{cov}(X_{t+m}, X_t)$ non dipende da t .

Nota 2.23. Costruiamo ora una funzione in R per impostare la visualizzazione dei grafici che faremo negli esempi successivi.

```

Fileps = function (file,larg,alt)
postscript(file,width=larg/2.54,height=alt/2.54,
horizontal=FALSE,onefile=FALSE,paper='special')

Impostagrafica = function (x1,x2,y1,y2,cornice=1,rxy=40,
assi=FALSE,marg=rep(0,4),file=NA,larg=9,alt=4,
par=c(bg='yellow'))
{if (!is.na(file)) Fileps(file=file,larg=larg,alt=alt)
par(mai=marg,lwd=1.2,cex=0.4); par(par)
plot(c(x1,x2),c(y1,y2),type='n',xlab='t',ylab='',asp=rxy,
axes=assi,frame.plot=as.logical(cornice))}

```

Nota 2.24. Le tre seguenti funzioni generano rispettivamente una serie temporale autoregressiva, un rumore bianco con deviazione standard uguale a 1 e una serie temporale a media mobile

```

autoregressiva1 = function (t1,t2,a,B,N=t1)
sapply(t1:t2,function (t) ruffini(B[(t-N):t],a))

biancounif = function (n,ds=1)
{B=runif(n); ds*(B-mean(B))/sd(B)}

mediamobile = function (t1,t2,theta,B)
{q=length(theta); theta=c(1,theta)
X=sapply(t1:t2,function (t) sum(B[t:(t-q)]*theta))}

```

Esempio 2.25. Con le istruzioni

```

n=100
B=biancounif(n,ds=0.1)

X=mediamobile(3,n,c(0.5,0.3),B)

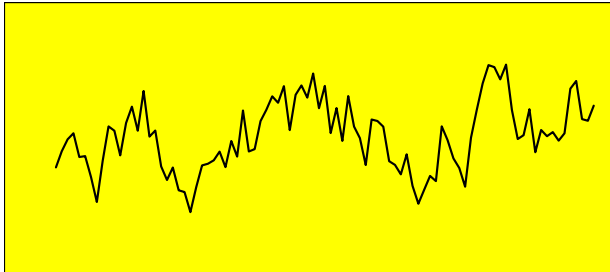
t1=8; Y=autoregressiva1(t1,n,0.7,B)

Impostagrafica(3,n,-0.3,0.3,file='0223-1.ps')
lines(3:n,X)
dev.off()

Impostagrafica(3,n,-0.3,0.3,file='0223-2.ps')
lines(t1:n,Y)
dev.off()

```

otteniamo ad esempio le figure



Queste serie temporali sono molto simili. Ciò non è un caso:
 La prima corrisponde a una serie temporale $\bigcirc_t X_t$ in cui

$$X_t = B_t + 0.5B_{t-1} + 0.3B_{t-2}$$

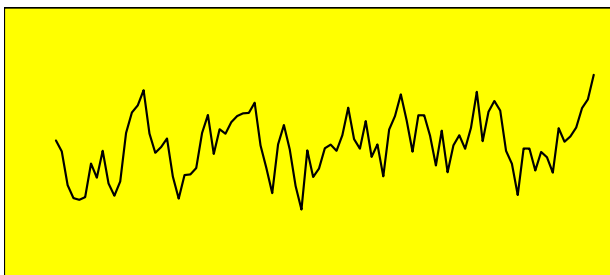
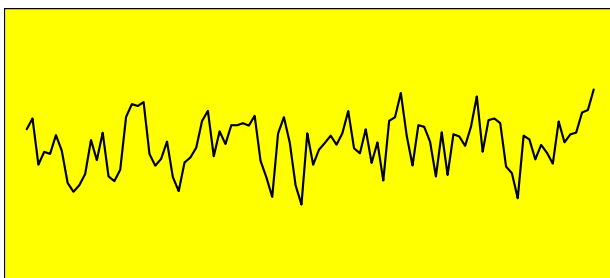
la seconda è una serie temporale $\bigcirc_t Y_t$ in cui

$$Y_t = B_t + 0.7B_{t-1} + 0.7^2B_{t-2} + 0.7^3B_{t-3} + \dots$$

dove si usa la stessa realizzazione dello stesso rumore bianco B , come approssimazione della sommatoria infinita $\tilde{Y} := \sum_{k=0}^{\infty} 0.7^k B_{t-k}$.

$\bigcirc_n Y_n$ è infatti di tipo MM $(0.7, 0.7^2, \dots, B)$, mentre una serie temporale di tipo AR (a, B) può essere detta di tipo MM (∞, B) , come vedremo.

Con le stesse istruzioni per un'altra realizzazione di B otteniamo le figure (di nuovo molto simili)



Nota 2.26. $\bigcirc_n X_n$ sia di tipo AR(a, B).

Per calcolare la funzione di autocovarianza di $\bigcirc_n X_n$ non abbiamo bisogno della rappresentazione esplicita data nel teorema 2.22. Infatti dall'oss. 2.19 sappiamo che $0 < |a| < 1$ e che $\gamma(0) = \frac{\sigma_B^2}{1 - |a|^2}$, mentre da $X_m = aX_{m-1} + B_m$ nel caso $m > 0$ otteniamo

$$\gamma(m) = \text{cov}(X_{t+m}, X_t) = \sigma_B^2 \frac{a^m}{1 - |a|^2} = a\sigma_B^2 \frac{a^{m-1}}{1 - |a|^2} = a\gamma(m-1)$$

Da questa relazione otteniamo direttamente la formula per $\gamma(m)$ del teorema 2.22. Le equazioni

$$\begin{aligned}\gamma(1) &= a\gamma(0) \\ \gamma(0) &= \frac{\sigma_B^2}{1 - |a|^2}\end{aligned}$$

sono dette *equazioni di Yule-Walker* della serie temporale data. Esse mostrano che le autocovarianze $\gamma(0)$ e $\gamma(1)$ da sole determinano i parametri a e σ_B^2 tramite

$$\begin{aligned}a &= \frac{\gamma(1)}{\gamma(0)} = \rho_1 \\ \sigma_B^2 &= \gamma(0) (1 - |\rho_1|^2)\end{aligned}$$

Definizione 2.27. Sia $p \in \mathbb{N} + 1$. La serie temporale $\bigcirc_n X_n$ si dice *autoregressiva di ordine p* , se sono soddisfatte le seguenti condizioni:

- (1) $\bigcirc_n X_n$ è stazionaria.
- (2) Esistono $a_1, \dots, a_p \in \mathbb{C}$ con $a_p \neq 0$ e un rumore bianco B tali che

$$X_t = a_1 X_{t-1} + \dots + a_p X_{t-p} + B_t$$

per ogni $t \in \mathbb{Z}$.

- (3) $\text{cov}(X_t, B_s) = 0$ per $t, s \in \mathbb{Z}$ con $s > t$.

Diremo allora che $\bigcirc_n X_n$ è di tipo AR(a_1, \dots, a_p, B).

Definizione 2.28. Siano $p, q \in \mathbb{N} + 1$ ed $a_1, \dots, a_p, \theta_1, \dots, \theta_q \in \mathbb{C}$ con $a_p \neq 0, \theta_q \neq 0$. La serie temporale $\bigcirc_n X_n$ si dice di tipo

ARMM($a_1, \dots, a_p; \theta_1, \dots, \theta_q; B$), se sono soddisfatte le seguenti condizioni:

- (1) $\bigcirc_n X_n$ è stazionaria.
- (2) $X_t = a_1 X_{t-1} + \dots + a_p X_{t-p} + B_t + \theta_1 B_{t-1} + \dots + \theta_q B_{t-q}$ per ogni $t \in \mathbb{Z}$
- (3) $\text{cov}(X_t, B_s) = 0$ per $t, s \in \mathbb{Z}$ con $s > t$.

Definizione 2.29. $\bigcirc_n X_n$ si chiama un *cammino casuale* (in inglese *random walk*), se sono soddisfatte le seguenti condizioni:

(1) $X_t = 0$ per $t < 0$.

(2) Esiste un rumore bianco B tale che

$$\begin{aligned} X_0 &= B_0 \text{ e} \\ X_t &= X_{t-1} + B_t \text{ per ogni } t \in \mathbb{N} + 1. \end{aligned}$$

Diremo allora che $\bigcirc_n X_n$ è il cammino casuale generato dal rumore bianco B .

Osservazione 2.30. $\bigcirc_n X_n$ sia il cammino casuale generato dal rumore bianco B . Allora:

(1) $X_t = \sum_{k=0}^t B_k$ per $t \geq 0$.

(2) $MX_t = 0$ per ogni $t \in \mathbb{Z}$.

(3) $\text{cov}(X_t, B_s) = 0$ per $0 \leq t < s$.

(4) $\text{cov}(X_{t+m}, X_t) = (t+1)\sigma_B^2$ per $t, m \in \mathbb{N}$.

Dall'ultima relazione vediamo in particolare che un cammino casuale non è stazionario.

Dimostrazione.

(1) Chiaro.

(2) Immediato da (1).

(3) Sia $0 \leq t < s$. Allora

$$\text{cov}(X_t, B_s) = \text{cov}\left(\sum_{k=0}^t B_k, B_s\right) = \sum_{k=0}^t \text{cov}(B_k, B_s) = 0.$$

(4) Abbiamo in primo luogo

$$\begin{aligned} \text{cov}(X_t, X_t) &= \text{cov}\left(\sum_{k=0}^t B_k, \sum_{j=0}^t B_j\right) = \sum_{k=0}^t \sum_{j=0}^t \text{cov}(B_k, B_j) \\ &= \sum_{k=0}^t \text{cov}(B_k, B_k) = (t+1)\sigma_B^2 \end{aligned}$$

Per $m > 0$ abbiamo inoltre

$$\begin{aligned} \text{cov}(X_{t+m}, X_t) &= \text{cov}\left(X_t + \sum_{k=t+1}^{t+m} B_k, X_t\right) \\ &= \text{cov}(X_t, X_t) + \sum_{k=t+1}^{t+m} \text{cov}(B_k, X_t) \stackrel{(3)}{=} \text{cov}(X_t, X_t) \end{aligned}$$

3. La misura spettrale

Definizione 3.1. Una funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ si dice *positivamente semidefinita*, se per ogni $r \in \mathbb{N} + 1$, ogni $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ ed ogni $n_1, \dots, n_r \in \mathbb{Z}$ si ha

$$\sum_{j=1}^r \sum_{k=1}^r \lambda_j \overline{\lambda_k} \varphi(n_j - n_k) \geq 0 \quad (*)$$

Chiediamo quindi in particolare che questa somma sia reale.

Se ogni volta che $(\lambda_1, \dots, \lambda_r) \neq 0$ la somma in (*) è > 0 , la funzione φ si dice *positivamente definita*.

Lemma 3.2. Siano $a, b \in \mathbb{C}$ tali che $a + b \in \mathbb{R}$, $a - b \in i\mathbb{R}$.

Allora $a = \overline{b}$.

Dimostrazione. Per ipotesi abbiamo $\text{Im}(a + b) = 0$ e $\text{Re}(a - b) = 0$, quindi $\text{Im}(a) = -\text{Im}(b)$ e $\text{Re}(a) = \text{Re}(b)$, e vediamo che $a = \overline{b}$.

Proposizione 3.3. La funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ sia positivamente semidefinita. Allora:

- (1) $\varphi(0) \geq 0$.
- (2) $\varphi(-m) = \overline{\varphi(m)}$ per ogni $m \in \mathbb{Z}$.
- (3) $|\varphi(m)| \leq \varphi(0)$ per ogni $m \in \mathbb{Z}$.

Dimostrazione. Seguiamo Kreiss/Neuhaus, pag 48-49.

(1) Ponendo $r = 1$, $\lambda_1 = 1$ e $n_1 = 0$ in (*) dalla def. 3.1 otteniamo $\varphi(0) \geq 0$.

(2) Per $r = 2$, $n_1 = m$, $n_2 = 0$ la (*) nella def. 3.1 diventa

$$\lambda_1 \overline{\lambda_1} \varphi(0) + \lambda_1 \overline{\lambda_2} \varphi(m) + \lambda_2 \overline{\lambda_1} \varphi(-m) + \lambda_2 \overline{\lambda_2} \varphi(0)$$

ovvero

$$(|\lambda_1|^2 + |\lambda_2|^2) \varphi(0) + \lambda_1 \overline{\lambda_2} \varphi(m) + \overline{\lambda_1} \lambda_2 \varphi(-m) \geq 0$$

Ponendo $\lambda_1 := \lambda_2 := 1$ si ha $2\varphi(0) + \varphi(m) + \varphi(-m) \geq 0$. Siccome $\varphi(0)$ è reale anche $\varphi(m) + \varphi(-m)$ deve essere reale.

Ponendo $\lambda_1 := 1$, $\lambda_2 := i$ si ottiene $2\varphi(0) - i(\varphi(m) - \varphi(-m)) \geq 0$, quindi $\varphi(m) - \varphi(-m) \in i\mathbb{R}$.

Dal lemma 3.2 otteniamo $\varphi(-m) = \overline{\varphi(m)}$.

(3) Utilizzando il punto (2) abbiamo quindi

$$(|\lambda_1|^2 + |\lambda_2|^2) \varphi(0) + 2 \text{Re} \lambda_1 \overline{\lambda_2} \varphi(m) \geq 0$$

(31) Consideriamo prima il caso $\varphi(0) = 0$. In questa ipotesi

$\text{Re} \lambda_1 \overline{\lambda_2} \varphi(m) \geq 0$ per ogni $\lambda_1, \lambda_2 \in \mathbb{C}$.

Per $\lambda_1 = 1$, $\lambda_2 = -\varphi(m)$ otteniamo $-\text{Re} |\varphi(m)|^2 \geq 0$ e quindi $\varphi(m) = 0$.

(32) Sia $\varphi(0) \neq 0$. Ponendo $a := \varphi(0)$, $b := \varphi(m)$ abbiamo

$$(|\lambda_1|^2 + |\lambda_2|^2) a + 2 \operatorname{Re} \lambda_1 \bar{\lambda}_2 b \geq 0 \text{ per ogni } \lambda_1, \lambda_2 \in \mathbb{C}.$$

Per ipotesi $a \neq 0$ e quindi $a > 0$ per il punto (1), cosicché anche

$$|\lambda_1|^2 + |\lambda_2|^2 + 2 \operatorname{Re} \frac{\lambda_1 \bar{\lambda}_2 b}{a} \geq 0 \text{ per ogni } \lambda_1, \lambda_2 \in \mathbb{C}.$$

Ponendo $\lambda_1 := 1, \lambda_2 := -b/a$ abbiamo perciò

$$1 + |b/a|^2 a - 2 \operatorname{Re} \frac{\bar{b}b}{\bar{a}a} \geq 0$$

ovvero $1 + |b/a|^2 - 2|b/a|^2 \geq 0$ e quindi $1 \geq |b/a|^2$

e questo implica $|b| \leq a$.

Osservazione 3.4. La funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{R}$ sia tale che $\varphi(m) = \varphi(-m)$ per ogni $m \in \mathbb{Z}$. Se per ogni $r \in \mathbb{N} + 1$, ogni $a_1, \dots, a_r \in \mathbb{R}$ ed ogni $n_1, \dots, n_r \in \mathbb{Z}$ si ha

$$\sum_{j=1}^r \sum_{k=1}^r a_j a_k \varphi(n_j - n_k) \geq 0$$

allora φ è positivamente semidefinita nel senso della def. 3.1, se consideriamo φ come funzione $\mathbb{Z} \rightarrow \mathbb{C}$.

Dimostrazione. Siano $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ ed $n_1, \dots, n_r \in \mathbb{Z}$ con $\lambda_k = a_k + ib_k$, e con $a_k, b_k \in \mathbb{R}$. Allora

$$(a_j + ib_j)(a_k - ib_k) = a_j a_k + b_j b_k + i(b_j a_k - a_j b_k)$$

Perciò l'ipotesi di simmetria implica che (tralasciando i limiti di somministrazione)

$$\begin{aligned} \sum_j \sum_k b_j a_k \varphi(n_j - n_k) &= \sum_j \sum_k b_j a_k \varphi(n_k - n_j) \\ &= \sum_j \sum_k a_j b_k \varphi(n_j - n_k) \end{aligned}$$

e vediamo che ci siamo ricondotti al caso reale.

Osservazione 3.5. Siano $X, Y \in \mathcal{L}^2$ ed $a, b \in \mathbb{C}$. Allora $\operatorname{cov}(X + a, Y + b) = \operatorname{cov}(X, Y)$.

Dimostrazione. Ciò è immediato dalla def. 1.10.

Proposizione 3.6. La funzione di autocovarianza di una serie temporale stazionaria è positivamente semidefinita.

Dimostrazione. $\bigcirc_n X_n$ sia una STS (μ, γ) .

Siano $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ ed $n_1, \dots, n_r \in \mathbb{Z}$. Allora

$$\begin{aligned} \sum_{j=1}^r \sum_{k=1}^r \lambda_j \bar{\lambda}_k \gamma(n_j - n_k) &= \sum_{j=1}^r \sum_{k=1}^r \lambda_j \bar{\lambda}_k \text{cov}(X_{n_j}, X_{n_k}) \\ &= \text{cov} \left(\sum_j \lambda_j X_{n_j}, \sum_k \lambda_k X_{n_k} \right) \\ &= \sigma^2 \left(\sum_j \lambda_j X_{n_j} \right) \geq 0 \end{aligned}$$

Teorema 3.7. $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ sia una funzione positivamente semidefinita. Allora esiste una serie temporale stazionaria $\bigcirc_n Y_n$ la cui funzione di autocovarianza coincide con φ .

Se φ è reale, anche la serie temporale $\bigcirc_n Y_n$ può essere scelta reale.

Dimostrazione. Ciò è una conseguenza non difficile del teorema di esistenza di Kolmogorov. Una dimostrazione si trova ad esempio in Brockwell/Davis, pag. 27.

Definizione 3.8. Sia $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ un'applicazione. Una *misura spettrale* per φ è una misura finita ν su Borel $((-\pi, \pi])$ tale per ogni $m \in \mathbb{Z}$ si abbia

$$\varphi(m) = \int_{(-\pi, \pi]} e^{imt} d\nu(t)$$

Nel seguito dimostreremo che la funzione φ possiede una misura spettrale se e solo se è positivamente semidefinita.

Proposizione 3.9. Se una funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ possiede una misura spettrale, allora φ è positivamente semidefinita.

Dimostrazione. Siano $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ e $n_1, \dots, n_r \in \mathbb{Z}$. Tralasciando di nuovo i limiti di sommazione abbiamo

$$\begin{aligned} \sum_j \sum_k \lambda_j \bar{\lambda}_k \varphi(n_j - n_k) &= \sum_j \sum_k \lambda_j \bar{\lambda}_k \int_{(-\pi, \pi]} e^{i(n_j - n_k)t} d\nu(t) \\ &= \int_{(-\pi, \pi]} \sum_j \sum_k \lambda_j \bar{\lambda}_k e^{i(n_j - n_k)t} d\nu(t) \\ &= \int_{(-\pi, \pi]} \left(\sum_j \lambda_j e^{in_j t} \sum_k \bar{\lambda}_k e^{in_k t} \right) d\nu(t) \\ &= \int_{(-\pi, \pi]} \left| \sum_k \lambda_k e^{in_k t} \right|^2 d\nu(t) \geq 0 \end{aligned}$$

Proposizione 3.10. *Se una funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ possiede una misura spettrale, questa è univocamente determinata.*

Dimostrazione. Siano ν_1, ν_2 due misure spettrali per φ .

Siccome $Borel((-\pi, \pi]) = \sigma\text{-algebra}(\{(a, b] \mid -\pi < a \leq b \leq \pi\})$, è sufficiente dimostrare che $\nu_1((a, b]) = \nu_2((a, b])$.

(1) Sia quindi $-\pi < a < b \leq \pi$.

È evidente che la funzione caratteristica $\chi_{(a,b]}$ può essere approssimata da una successione monotona decrescente $\bigcirc_n f_n$ di funzioni continue (che possiamo scegliere lineari a tratti) $f_n : [-\pi, \pi] \rightarrow [0, 1]$ e tali che $f(-\pi) = f(\pi)$. Per il teorema di Levi sulla convergenza monotona abbiamo allora

$$\nu_k((a, b]) = \int_{(-\pi, \pi]} \chi_{(a,b]} d\nu_k(x) = \lim_{n \rightarrow \infty} \int_{(-\pi, \pi]} f_n(x) d\nu_k(x)$$

per $k = 1, 2$.

(2) È quindi sufficiente dimostrare che

$$\int_{(-\pi, \pi]} f(x) d\nu_1(x) = \int_{(-\pi, \pi]} f(x) d\nu_2(x)$$

per ogni funzione continua $f_n : [-\pi, \pi] \rightarrow [0, 1]$ con $f(-\pi) = f(\pi)$.

Per il teorema di approssimazione di Weierstrass ogni tale funzione può essere approssimata uniformemente tramite polinomi trigonometrici della forma $\bigcirc_t \sum_{m=-N}^N a_m e^{imt}$ e quindi è sufficiente dimostrare che

$$\int_{(-\pi, \pi]} e^{imt} d\nu_1(t) = \int_{(-\pi, \pi]} e^{imt} d\nu_2(t)$$

per ogni $m \in \mathbb{Z}$. Ma per ipotesi i due integrali sono uguali a $\varphi(m)$.

Definizione 3.11. Per uno spazio topologico X denotiamo con $\mathcal{M}(X)$ l'insieme delle misure finite su $Borel(X)$.

$C_b(X, \mathbb{R})$ sia, come d'uso, l'insieme delle funzioni reali continue e limitate definite su X .

Definizione 3.12. Sia X uno spazio topologico. Un insieme $\mathcal{N} \subset \mathcal{M}(X)$ si dice *concentrato sui compatti* (in inglese *tight*), se per ogni $\varepsilon > 0$ esiste un sottoinsieme compatto $K \subset X$ tale che $\nu(X \setminus K) < \varepsilon$ per ogni $\nu \in \mathcal{N}$.

\mathcal{N} si dice *limitato*, se l'insieme $\{\nu(X) \mid \nu \in \mathcal{N}\}$ è limitato.

Definizione 3.13. X sia uno spazio topologico, $\theta \in \mathcal{M}(X)$ e $\bigcirc_n \nu_n$ una successione di elementi di $\mathcal{M}(X)$. Diciamo che la successione $\bigcirc_n \nu_n$ converge *debolmente* a θ e scriviamo $\bigcirc_n \nu_n \xrightarrow{\text{deb}} \theta$, se per ogni $f \in C_b(X, \mathbb{R})$ si ha

$$\bigcirc_n \int f(X) d\nu_n(X) \rightarrow \int f(X) d\theta(X)$$

Definizione 3.14. Sia X uno spazio topologico. Un insieme $\mathcal{N} \subset \mathcal{M}(X)$ si dice *relativamente sequenzialmente compatto*, se per ogni successione $\bigcirc_n \nu_n$ di elementi di \mathcal{N} esistono una sottosuccessione $\bigcirc_k \nu_{n_k}$ e un elemento $\theta \in \mathcal{M}(X)$ tali che $\bigcirc_k \nu_{n_k} \xrightarrow{\text{deb}} \theta$.

Teorema 3.15 (teorema di Prokhorov). X sia uno spazio metrico completo ed $\mathcal{N} \subset \mathcal{M}(X)$. Allora sono equivalenti:

- (1) \mathcal{N} è relativamente sequenzialmente compatto.
- (2) \mathcal{N} è concentrato sui compatti e limitato.

Dimostrazione. La dimostrazione di questo teorema è molto complicata e si trova ad esempio in Elstrodt, pagg. 392-398, e Billingsley, pagg. 57-64.

Lemma 3.16. Siano $g : \mathbb{Z} \rightarrow \mathbb{C}$ una funzione ed $n \in \mathbb{N} + 1$. Allora

$$\sum_{k=-(n-1)}^{n-1} (n - |k|) g(k) = \sum_{j=1}^n \sum_{k=1}^n g(j - k)$$

Dimostrazione. Questa formula si trova in Kreiss/Neuhaus, pag. 364. La dimostriamo per induzione su n .

$n = 1$: Sia a sinistra che a destra abbiamo soltanto $g(0)$.

$n \rightarrow n + 1$:

$$\begin{aligned} \sum_{k=-n}^n (n + 1 - |k|) g(k) &= \sum_{k=-n}^n g(k) + \sum_{k=-n}^n (n - |k|) g(k) \\ &= \sum_{k=-n}^n g(k) + \sum_{k=-(n-1)}^{n-1} (n - |k|) g(k) \\ &\stackrel{\text{ind}}{=} \sum_{k=-n}^n g(k) + \sum_{j=1}^n \sum_{k=1}^n g(j - k) \\ &= \sum_{k=-n}^n g(k) + \sum_{j=1}^{n+1} \sum_{k=1}^n g(j - k) \\ &\quad - \sum_{k=1}^n g(n + 1 - k) \\ &= \sum_{k=-n}^n g(k) + \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} g(j - k) \\ &\quad - \sum_{k=1}^n g(n + 1 - k) - \sum_{j=1}^{n+1} g(j - n - 1) \end{aligned}$$

Però $\sum_{k=1}^n g(n + 1 - k) + \sum_{j=1}^{n+1} g(j - n - 1) = \sum_{k=-n}^n g(k)$.

Lemma 3.17. $\varphi : \mathbb{Z} \longrightarrow \mathbb{C}$ sia una funzione positivamente semidefinita. Per ogni $n \in \mathbb{N} + 1$ definiamo una funzione $f_n : [-\pi, \pi] \longrightarrow \mathbb{C}$ tramite

$$f_n(t) := \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \varphi(k) e^{-ikt}$$

Per ogni $n \in \mathbb{N} + 1$ e per ogni $t \in [-\pi, \pi]$ allora:

$$(1) \quad f_n(t) = \frac{1}{2\pi n} \sum_{j=1}^n \sum_{k=1}^n \varphi(j-k) e^{-ijt} e^{ikt}.$$

(2) $f_n(t)$ è reale e ≥ 0 .

(3) $f_n \in C_b([-\pi, \pi], \mathbb{R})$.

(4) Per ogni $m \in \mathbb{Z}$ tale che $n > |m|$ vale

$$\int_{-\pi}^{\pi} f_n(t) e^{imt} dt = \left(1 - \frac{|m|}{n}\right) \varphi(m)$$

$$(5) \quad \int_{-\pi}^{\pi} f_n(t) dt = \varphi(0).$$

Dimostrazione. (1) Ciò segue dal lemma 3.16.

(2) Ciò segue dal punto (1), usando l'ipotesi che f sia positivamente semidefinita.

(3) Chiaro.

(4)

$$\begin{aligned} \int_{-\pi}^{\pi} f_n(t) e^{imt} dt &= \frac{1}{2\pi} \sum_{k=-(n-1)}^{n-1} \left(1 - \frac{|k|}{n}\right) \varphi(k) \int_{-\pi}^{\pi} e^{i(m-k)t} dt \\ &= \frac{1}{2\pi} \left(1 - \frac{|m|}{n}\right) \varphi(m) 2\pi = \left(1 - \frac{|m|}{n}\right) \varphi(m) \end{aligned}$$

perché, come noto,

$$\int_{-\pi}^{\pi} e^{i(m-k)t} dt = \begin{cases} 2\pi & \text{per } m = k \\ 0 & \text{altrimenti} \end{cases}$$

(5) Caso particolare del punto (4) per $m = 0$.

Teorema 3.18 (teorema di Herglotz). Una funzione $\varphi : \mathbb{Z} \longrightarrow \mathbb{C}$ possiede una misura spettrale se e solo se è positivamente semidefinita.

Dimostrazione. (1) Se φ possiede una misura spettrale allora è positivamente semidefinita per la prop. 3.9.

(2) Sia viceversa φ positivamente semidefinita.

Per $B \in \text{Borel}([-\pi, \pi])$ ed $n \in \mathbb{N} + 1$ sia $\nu_n(B) := \int_B f_n(t) dt$, dove l'integrale è calcolato nel senso di Lebesgue.

Per il punto (5) del lemma 3.17 $\nu_n \in \mathcal{M}([-\pi, \pi])$; per la stessa ragione l'insieme $\mathcal{N} := \{\nu_n \mid n \in \mathbb{N} + 1\}$ è limitato.

Siccome $[-\pi, \pi]$ è compatto, \mathcal{N} è inoltre sicuramente concentrato sui

compatti. Per il teorema 3.15 esistono una sottosuccessione

$\bigcirc_k \nu_{n_k}$ e un elemento $\theta \in \mathcal{M}([-\pi, \pi])$ tali che $\bigcirc_k \nu_{n_k} \xrightarrow{\text{deb}} \theta$.

Sia $m \in \mathbb{Z}$. Siccome $\bigcirc_k n_k \rightarrow \infty$, dal punto (4) del lemma 3.17 vediam

mo che $\lim_{k \rightarrow \infty} \int_{-\pi}^{\pi} f_{n_k}(t) e^{imt} dt = \varphi(m)$.

Per la formula di trasformazione degli integrali nota dall'analisi però

$$\int_{-\pi}^{\pi} f_n(t) e^{imt} dt = \int_{[-\pi, \pi]} e^{imt} d\nu_n t$$

cosicché, tenendo conto del punto (3) del lemma 3.17 e della def. 3.13,

$$\begin{aligned} \varphi(m) &= \lim_{k \rightarrow \infty} \int_{-\pi}^{\pi} f_{n_k}(t) e^{imt} d\nu_{n_k}(t) = \lim_{k \rightarrow \infty} \int_{[-\pi, \pi]} e^{imt} d\nu_{n_k}(t) \\ &= \int_{-\pi}^{\pi} e^{imt} d\theta(t) = \int_{(-\pi, \pi]} e^{imt} d\theta(t) + e^{-i\pi m} \theta(\{-\pi\}) \\ &= \int_{(-\pi, \pi]} e^{imt} d\nu(t) \end{aligned}$$

se per $B \in \text{Borel}((-\pi, \pi])$ poniamo

$$\nu(B) := \begin{cases} \theta(B) & \text{se } \pi \notin B \\ \theta(B) + \theta(\{-\pi\}) & \text{se } \pi \in B \end{cases}$$

ν è allora una misura spettrale per φ .

Abbiamo seguito Kreiss/Neuhaus, pagg. 49-51. Un'altra dimostrazione si trova in Katznelson, pagg. 38-39. La dimostrazione in Brockwell/Davis, pagg. 117-119, è molto simile alla nostra.

L'analogo del teorema di Herglotz per processi stocastici a tempo continuo è il teorema di Wiener-Khinchin, al quale il teorema di Herglotz può essere ricondotto; cfr. Priestley, pagg. 218-224.

Corollario 3.19. Per una funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ sono equivalenti:

- (1) φ è la funzione di autocovarianza di una serie temporale stazionaria.
- (2) φ è positivamente semidefinita.
- (3) φ possiede una misura spettrale.

Definizione 3.20. La misura spettrale di una serie temporale stazionaria è la misura spettrale della sua funzione di autocovarianza.

Questa definizione è giustificata in virtù del corollario 3.19.

Definizione 3.21. Una serie temporale $\bigcirc_n X_n$ si dice *armonica*, se esistono numeri $\omega_1, \dots, \omega_N \in (0, \pi]$, variabili aleatorie reali $A_1, \dots, A_N, B_1, \dots, B_N$ tali che siano soddisfatte le seguenti condizioni:

- (1) $X_t = \sum_{k=1}^N (A_k \cos \omega_k t + B_k \sin \omega_k t)$ per ogni $t \in \mathbb{Z}$.

- (2) $A_k, B_k \in \mathcal{L}^2$ per ogni k .
- (3) $MA_k = MB_k = 0$ per ogni k .
- (4) $\text{cov}(A_k, B_k) = 0$ per ogni k .
- (5) $\text{cov}(A_j, A_k) = \text{cov}(B_j, B_k) = 0$ per $j \neq k$.
- (6) $\sigma^2(A_k) = \sigma(B_k) =: \sigma_k^2$ per ogni k .

Diremo allora che $\bigcirc_n X_n$ è una STA $(\omega_1, \dots, \omega_N, \sigma_1^2, \dots, \sigma_N^2)$

Proposizione 3.22. $\bigcirc_n X_n$ sia una STA $(\omega_1, \dots, \omega_N, \sigma_1^2, \dots, \sigma_N^2)$.

Allora $\bigcirc_n X_n$ è una STS $(0, \gamma)$ con

$$\gamma(m) = \sum_{k=1}^N \sigma_k^2 \cos \omega_k m$$

per ogni $m \in \mathbb{Z}$.

Dimostrazione. (1) È chiaro che $X_t \in \mathcal{L}^2$ (cfr. oss. 1.9).

(2) È immediato anche che $MX_t = 0$ per ogni $t \in \mathbb{Z}$.

(3) Siano $t, m \in \mathbb{Z}$. Con la notazione della def. 3.21 abbiamo

$$\begin{aligned} X_{t+m} &= \sum_{k=1}^N (A_k \cos \omega_k (t+m) + B_k \sin \omega_k (t+m)) \\ &= \sum_{k=1}^N A_k (\cos \omega_k t \cos \omega_k m - \sin \omega_k t \sin \omega_k m) + \\ &\quad + \sum_{k=1}^N B_k (\sin \omega_k t \cos \omega_k m + \sin \omega_k m \cos \omega_k t) \\ X_t &= \sum_{k=1}^N A_k \cos \omega_k t + \sum_{k=1}^N B_k \sin \omega_k t \end{aligned}$$

cosicché dalla prop 1.14 e dalle condizioni di ortogonalità della def. 3.21 otteniamo

$$\begin{aligned} \text{cov}(X_{t+m}, X_t) &= \sum_{k=1}^N \sigma_k^2 (\cos^2 \omega_k t \cos \omega_k m - \sin \omega_k t \sin \omega_k m \cos \omega_k t) \\ &\quad + \sum_{k=1}^N \sigma_k^2 (\sin^2 \omega_k t \cos \omega_k m + \sin \omega_k m \cos \omega_k t \sin \omega_k t) \\ &= \sum_{k=1}^N \sigma_k^2 \cos \omega_k m (\cos^2 \omega_k t + \sin^2 \omega_k t) \\ &= \sum_{k=1}^N \sigma_k^2 \cos \omega_k m \end{aligned}$$

(4) Ciò mostra la stazionarietà di $\bigcirc_n X_n$ e in particolare che la funzione di autocovarianza si calcola come nell'enunciato.

Proposizione 3.23. $\bigcirc_n X_n$ sia una STA $(\omega_1, \dots, \omega_N, \sigma_1^2, \dots, \sigma_N^2)$. Allora la misura spettrale di $\bigcirc_n X_n$ è data da

$$\nu(B) := \sum_{k=1}^N \lambda_k \sigma_k^2$$

con

$$\lambda_k := \begin{cases} \frac{\chi_B(\omega_k) + \chi_B(-\omega_k)}{2} & \text{per } 0 < \omega_k < \pi \\ \chi_B(\pi) & \text{per } \omega_k = \pi \end{cases}$$

Dimostrazione. Ciò segue facilmente dalla prop. 3.22 tenendo conto della relazione

$$\cos \omega_k m = \frac{e^{i\omega_k m} + e^{-i\omega_k m}}{2}$$

Lemma 3.24 (teorema della convergenza maggiorata). $(\Omega', \mathcal{B}, \mu)$ sia uno spazio di misura e $g, f, f_1, f_2, \dots : \Omega \rightarrow \mathbb{C}$ funzioni misurabili tali che siano soddisfatte le seguenti condizioni:

- (1) $\bigcirc_n f_n \rightarrow f$ quasi dappertutto rispetto a μ .
- (2) $g \in \mathcal{L}^1(\mu)$.
- (3) $|f_n| \leq |g|$ quasi dappertutto rispetto a μ .

Allora:

$$(A) \lim_{n \rightarrow \infty} \int |f_n(x) - f(x)| d\mu(x) = 0.$$

$$(B) \lim_{n \rightarrow \infty} \int f_n(x) dx = \int f(x) dx.$$

Dimostrazione. Corsi di analisi. Si noti che (B) segue da (A) per la disuguaglianza $\left| \int h(x) dx \right| \leq \int |h(x)| dx$ valida per ogni funzione μ -integrabile h .

Osservazione 3.25. Per $A \subset \mathbb{Z}$ sia $\mu := |A|$, quindi $\mu(A) = \infty$ se A non è finito. Allora μ è una misura su $\mathcal{P}(\mathbb{Z})$ che verrà detta *misura di conteggio*. Per $f \in \mathbb{C}^{\mathbb{Z}}$ si ha $f \in \mathcal{L}^1(\mu) \iff \sum_{n=-\infty}^{\infty} |f(n)| < \infty$ e in tal caso

$$\int f(k) d\mu(k) = \sum_{n=-\infty}^{\infty} f(n).$$

Cfr. Elstrodt, 139.

Lemma 3.26 (Kronecker). Sia $\bigcirc_n c_n \in \mathbb{Z}$ tale che $\sum_{n=-\infty}^{\infty} |c_n| < \infty$.

Allora $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=-n}^n |kc_k| = 0$.

Dimostrazione. Come nell'oss. 3.25 consideriamo lo spazio di misura $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}), \mu)$, dove μ è la misura di conteggio.

Con $g := \bigcirc_n c_n$, $f := 0$ ed

$$f_n(k) := \begin{cases} \frac{|kc_k|}{n} & \text{per } |k| \leq n \\ 0 & \text{altrimenti} \end{cases}$$

siamo nelle ipotesi del lemma 3.24, perché $\lim_{n \rightarrow \infty} f_n(k) = 0$ per ogni $k \in \mathbb{Z}$.

$$\text{Perciò } \lim_{n \rightarrow \infty} \int f_n(k) d\mu(k) = 0.$$

$$\text{Ma } \int f_n(k) d\mu(k) = \sum_{k=-\infty}^{\infty} f_n(k) = \frac{1}{n} \sum_{k=-n}^n |kc_k|$$

Teorema 3.27 (formula di inversione). La funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ sia positivamente semidefinita e tale che $\sum_{m=-\infty}^{\infty} |\varphi(m)| < \infty$.

$$\text{Per } t \in [-\pi, \pi] \text{ sia } f(t) := \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \varphi(m) e^{-imt}.$$

Allora

(1) $f(t)$ è reale e ≥ 0 per ogni $t \in [-\pi, \pi]$.

(2) La funzione $f : [-\pi, \pi] \rightarrow \mathbb{R}$ è continua.

$$(3) \varphi(m) = \int_{-\pi}^{\pi} e^{imt} f(t) dt \quad \text{per ogni } m \in \mathbb{Z}.$$

$$(4) \text{ In particolare } \varphi(0) = \int_{-\pi}^{\pi} f(t) dt.$$

La misura spettrale di φ possiede quindi la densità spettrale continua f . Questa è naturalmente univocamente determinata.

Dimostrazione. (1) L'ipotesi che $\sum_{m=-\infty}^{\infty} |\varphi(m)| < \infty$ implica che le somme parziali $\sum_{|m| \leq n} \varphi(m) e^{-imt}$ convergono uniformemente ad f , quindi f è continua.

(2) Per $n \in \mathbb{N} + 1$ definiamo una funzione $f_n : [-\pi, \pi] \rightarrow \mathbb{C}$ tramite

$$f_n(t) := \frac{1}{2\pi} \sum_{|k| < n} \left(1 - \frac{|k|}{n}\right) \varphi(k) e^{-ikt}$$

come nel lemma 3.17. Per ogni $t \in [-\pi, \pi]$ allora

$$\begin{aligned}
|f(t) - f_n(t)| &= \left| \frac{1}{2\pi} \sum_{|k| \leq n} \varphi(k) e^{-ikt} + \frac{1}{2\pi} \sum_{|k| > n} \varphi(k) e^{-ikt} \right. \\
&\quad \left. - \frac{1}{2\pi} \sum_{|k| < n} \left(1 - \frac{|k|}{n}\right) \varphi(k) e^{-ikt} \right| \\
&= \left| \frac{1}{2\pi} \sum_{|k| \geq n} \varphi(k) e^{-ikt} + \frac{1}{2\pi n} \sum_{|k| < n} |k| \varphi(k) e^{-ikt} \right| \\
&\leq \frac{1}{2\pi} \sum_{|k| \geq n} |\varphi(k)| + \frac{1}{2\pi n} \sum_{|k| < n} |k| |\varphi(k)|
\end{aligned}$$

Il primo sommando tende a 0 perché $\sum_{m=-\infty}^{\infty} |\varphi(m)| < \infty$, mentre il secondo sommando tende a 0 per il lemma 3.26.

(3) Perciò $\lim_{n \rightarrow \infty} f_n(t) = f(t)$ per ogni $t \in [-\pi, \pi]$ e dal lemma 3.27 segue che f è reale e ≥ 0 .

(4) Rimane da dimostrare che $\varphi(m) = \int_{-\pi}^{\pi} e^{imt} f_n(t) dt$

per ogni $m \in \mathbb{Z}$.

Siccome però $0 \leq f_n(t) \leq |f(t)|$ per ogni $t \in [-\pi, \pi]$ ed ogni n , ciò segue dal lemma 3.24.

Corollario 3.28. *La funzione $\varphi : \mathbb{Z} \rightarrow \mathbb{C}$ sia positivamente semidefinita, reale e tale che $\sum_{m=-\infty}^{\infty} |\varphi(m)| < \infty$. Sia f la densità spettrale continua di φ come nel teorema 3.27. Per ogni $m \in \mathbb{Z}$ allora*

$$\varphi(m) = \int_{-\pi}^{\pi} f(t) \cos mtdt = 2 \int_0^{\pi} f(t) \cos mtdt$$

Dimostrazione. Per la prop. 3.3 abbiamo $\varphi(-m) = \varphi(m)$ e quindi

$$\varphi(m) = \int_{-\pi}^{\pi} e^{imt} f(t) dt = \int_{-\pi}^{\pi} e^{-imt} f(t) dt$$

cosicché

$$2\varphi(m) = \int_{-\pi}^{\pi} (e^{imt} + e^{-imt}) f(t) dt = 2 \int_{-\pi}^{\pi} f(t) \cos mtdt$$

Siccome nelle nostre ipotesi

$$\begin{aligned}
f(-t) &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \varphi(m) e^{imt} = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \varphi(-m) e^{-imt} \\
&= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \varphi(m) e^{-imt} = f(t)
\end{aligned}$$

abbiamo anche $\int_{-\pi}^{\pi} f(t) \cos mtdt = 2 \int_{-\pi}^{\pi} f(t) \cos mt.$

Definizione 3.29. $\bigcirc_n X_n$ sia una STS (μ, γ) tale che $\sum_{m=-\infty}^{\infty} |\gamma(m)| < \infty$. Allora la densità spettrale continua di γ si chiama la densità spettrale di $\bigcirc_n X_n$.

Esempio 3.30. $\bigcirc_n X_n$ sia un RB $(0, \sigma^2)$. La funzione di autocovarianza è data da

$$\gamma(m) = \begin{cases} 0 & \text{per } m \neq 0 \\ \sigma^2 & \text{per } m = 0 \end{cases}$$

come sappiamo dall'esempio 2.12.

La densità spettrale è quindi costante e uguale ad $f(t) = \frac{\sigma^2}{2\pi}$ per ogni $t \in [-\pi, \pi]$.

Esempio 3.31. $\bigcirc_n X_n$ sia un MM (θ, B) con $\theta \in \mathbb{R}$. La funzione di autocovarianza è data da

$$\gamma(m) = \begin{cases} (1 + \theta^2) \sigma_B^2 & \text{per } m = 0 \\ \theta \cdot \sigma_B^2 & \text{per } m = \pm 1 \\ 0 & \text{altrimenti} \end{cases}$$

come sappiamo dal cor. 2.16.

La densità spettrale è quindi data da

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \sum_{m=-1}^1 \gamma(m) e^{-imt} = \frac{\sigma_B^2}{\pi} \left(\frac{1 + \theta^2}{2} + \frac{\theta e^{it} + \theta e^{-it}}{2} \right) \\ &= \frac{\sigma_B^2}{2\pi} (1 + \theta^2 + 2\theta \cos t) \end{aligned}$$

Esempio 3.32. $\bigcirc_n X_n$ sia un MM (θ_1, θ_2, B) con $\theta_1, \theta_2 \in \mathbb{R}$. In questo caso la funzione di autocovarianza è data da

$$\begin{aligned} \gamma(0) &= (1 + \theta_0^2 + \theta_1^2) \sigma_B^2 \\ \gamma(1) &= (\theta_1 + \theta_2 \theta_1) \sigma_B^2 = \gamma(-1) \\ \gamma(2) &= \theta_2 \sigma_B^2 = \gamma(-2) \\ \gamma(m) &= 0 \text{ per } |m| \geq 3. \end{aligned}$$

come noto dalla prop. 2.14.

Perciò la densità spettrale è data da

$$\begin{aligned}
 f(t) &= \frac{1}{2\pi} \sum_{m=-2}^2 \gamma(m) e^{-imt} \\
 &= \frac{1}{2\pi} (\gamma(0) + \gamma(1)(e^{it} + e^{-it}) + \gamma(2)(e^{2it} + e^{-2it})) \\
 &= \frac{1}{2\pi} (\gamma(0) + 2\gamma(1) \cos t + 2\gamma(2) \cos 2t) \\
 &= \frac{\sigma_B^2}{2\pi} (1 + \theta_0^2 + \theta_1^2 + 2(\theta_1 + \theta_2\theta_1) \cos t + 2\theta_2 \cos 2t)
 \end{aligned}$$

Esempio 3.33. \bigcirc X_n sia un AR (a, B) con $a \in \mathbb{C}$ e $0 < |a| < 1$ definita come nel teorema 2.22. In questo caso la funzione di autocovarianza è data da

$$\gamma(m) = \sigma_B^2 \frac{a^{|m|}}{1 - |a|^2} \quad \text{per } m \in \mathbb{Z}.$$

La densità spettrale è quindi data da

$$\begin{aligned}
 f(t) &= \frac{1}{2\pi} \sigma_B^2 \sum_{m=-\infty}^{\infty} \frac{a^{|m|}}{1 - |a|^2} e^{-imt} \\
 &= \frac{\sigma_B^2}{2\pi(1 - |a|^2)} \left\{ 1 + \sum_{m=1}^{\infty} (ae^{it})^m + \sum_{m=1}^{\infty} (ae^{-it})^m \right\} \\
 &= \frac{\sigma_B^2}{2\pi(1 - |a|^2)} \left(\frac{1}{1 - ae^{it}} + \frac{1}{1 - ae^{-it}} - 1 \right) \\
 &= \frac{\sigma_B^2}{2\pi(1 - |a|^2)} \left(\frac{1 - a^2}{1 + a^2 - 2a \cos t} \right)
 \end{aligned}$$

4. Spazi vettoriali con un prodotto scalare

Situazione 4.1. Sia $\mathbb{K} := \mathbb{R}$ oppure \mathbb{C} .

V sia uno spazio vettoriale su \mathbb{K} e $\|\cdot\|$ un prodotto scalare su V .

Osservazione 4.2. Per definizione $\|\cdot\|$ è un'applicazione $V \times V \rightarrow \mathbb{K}$ con le seguenti proprietà, valide per ogni $x, y, z \in V$:

- (1) $\|x + y, z\| = \|x, z\| + \|y, z\|$.
- (2) $\|\lambda x, y\| = \lambda \|x, y\|$ per ogni $\lambda \in \mathbb{K}$.
- (3) $\|y, x\| = \overline{\|x, y\|}$.

Nel caso reale sia ha quindi $\|y, x\| = \|x, y\|$.

- (4) $\|x, x\| \geq 0$.
- (5) $\|x, x\| = 0 \Rightarrow x = 0$.

Si noti che la condizione (2) implica $\|0, 0\| = 0$.

Proposizione 4.3 (disuguaglianza di Cauchy-Schwarz).

Siano $x, y \in V$. Allora $\|x, y\|^2 = \|x, y\| \|y, x\| \leq \|x, x\| \|y, y\|$.

Dimostrazione. (1) $\|x, y\|^2 = \|x, y\| \overline{\|x, y\|} = \|x, y\| \|y, x\|$.

(2) Rimane da dimostrare la disuguaglianza. Per ogni $\lambda \in \mathbb{K}$ abbiamo

$$0 \leq \|x + \lambda y, x + \lambda y\| = \|x, x\| + \bar{\lambda} \|x, y\| + \lambda \|y, x\| + \lambda \bar{\lambda} \|y, y\|$$

La disuguaglianza nell'enunciato è banalmente vera per $y = 0$. Possiamo quindi assumere che $y \neq 0$ e porre $\lambda := -\frac{\|x, y\|}{\|y, y\|}$, ottenendo

$$\|x, x\| - \frac{\|x, y\| \|y, x\|}{\|y, y\|} - \frac{\|y, x\| \|x, y\|}{\|y, y\|} + \frac{\|x, y\| \|y, x\|}{\|y, y\|} \geq 0$$

ovvero

$$0 \leq \|x, x\| \|y, y\| - \|x, y\| \|y, x\| = \|x, x\| \|y, y\| - \|x, y\|^2$$

Osservazione 4.4. Per $x, y \in V$ sono equivalenti:

- (1) x ed y sono linearmente indipendenti.
- (2) $\|x, y\|^2 < \|x, x\| \|y, y\|$.

Dimostrazione. (1) \Rightarrow (2) : x ed y siano linearmente indipendenti. Per ogni $\lambda \in \mathbb{K}$ allora $x + \lambda y \neq 0$, per cui nella dimostrazione della prop. 4.3 possiamo sostituire il segno di disuguaglianza \leq con $<$.

(2) \Rightarrow (1) : Sia ad esempio $y = \lambda x$ con $\lambda \in \mathbb{K}$. Allora

$$\begin{aligned} \|x, y\| \|y, x\| &= \|x, \lambda x\| \|\lambda x, x\| = \lambda \bar{\lambda} \|x, x\|^2 \\ \|x, x\| \|y, y\| &= \|x, x\| \|\lambda x, \lambda x\| = \lambda \bar{\lambda} \|x, x\|^2 \end{aligned}$$

Definizione 4.5. Per $x \in V$ poniamo $\|x\| := \sqrt{\|x, x\|}$.

Osservazione 4.6. Siano $x, y \in V$. Allora

- (1) $\|x, y\| \leq \|x\| \|y\|$.
- (2) $\|x, y\| + \|y, x\| \leq 2\|x\| \|y\|$.

Dimostrazione. (1) Questa è una riformulazione della prop. 4.3.

$$(2) \|x, y\| + \|y, x\| = 2 \operatorname{Re} \|x, y\| \leq 2 \|x, y\| \leq 2\|x\| \|y\|.$$

Lemma 4.7. L'applicazione $\| \cdot \| : V \rightarrow \mathbb{R}$ è una norma su V , sono cioè soddisfatte le seguenti condizioni, per ogni $x, y \in V$:

- (1) $\|x\| \geq 0$.
- (2) $\|x\| = 0 \Rightarrow x = 0$.
- (3) $\|\lambda x\| = |\lambda| \|x\|$ per ogni $\lambda \in \mathbb{K}$.
- (4) $\|x + y\| \leq \|x\| + \|y\|$.

Si noti che la condizione (2) implica $\|0\| = 0$.

Dimostrazione. (1) Per definizione.

(2) Segue dall'oss. 4.2.

$$(3) \|\lambda x\| = \sqrt{\|\lambda x, \lambda x\|} = \sqrt{\lambda \bar{\lambda} \|x, x\|} = \sqrt{|\lambda|^2 \|x\|^2} = |\lambda| \|x\|.$$

(4) Utilizzando l'oss. 4.6 abbiamo

$$\begin{aligned} \|x + y\|^2 &= \|x + y, x + y\| = \|x, x\| + \|x, y\| + \|y, x\| + \|y, y\| \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

Osservazione 4.8. Se per $x, y \in V$ poniamo $d(x, y) := \|x - y\|$, allora (V, d) diventa uno spazio metrico. Concetti topologici si riferiranno sempre alla metrica d .

Proposizione 4.9 (legge del parallelogramma). Siano $x, y \in V$. Allora

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

Dimostrazione. Come nella dimostrazione del lemma 4.7 abbiamo

$$\begin{aligned} \|x + y\|^2 &= \|x + y, x + y\| = \|x, x\| + \|x, y\| + \|y, x\| + \|y, y\| \\ \|x - y\|^2 &= \|x - y, x - y\| = \|x, x\| - \|x, y\| - \|y, x\| + \|y, y\| \end{aligned}$$

Sommando le due relazioni otteniamo

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x, x\| + 2\|y, y\| = 2\|x\|^2 + 2\|y\|^2$$

Osservazione 4.10. (1) Sia $\mathbb{K} = \mathbb{R}$. Allora

$$4\|x, y\| = \|x + y\|^2 - \|x - y\|^2$$

per ogni $x, y \in V$.

(2) Sia $\mathbb{K} = \mathbb{C}$. Allora

$$4\|x, y\| = \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2$$

per ogni $x, y \in V$.

In entrambi i casi siamo quindi in grado di esprimere il prodotto scalare tramite la norma.

Dimostrazione. Siano $x, y \in V$.

(1) Sia $\mathbb{K} = \mathbb{R}$. Allora il prodotto scalare è simmetrico e si ha

$$\begin{aligned} \|x + y\|^2 - \|x - y\|^2 &= \|x + y, x + y\| - \|x - y, x - y\| \\ &= \|x, x\| + 2\|x, y\| + \|y, y\| \\ &\quad - \|x, x\| + 2\|x, y\| - \|y, y\| \\ &= 4\|x, y\| \end{aligned}$$

(2) Sia $\mathbb{K} = \mathbb{C}$. Come nella dimostrazione della prop. 4.9

$$\begin{aligned} \|x + y\|^2 - \|x - y\|^2 &= \|x + y, x + y\| - \|x - y, x - y\| \\ &= \|x, x\| + \|x, y\| + \|y, x\| + \|y, y\| \\ &\quad - \|x, x\| + \|x, y\| + \|y, x\| - \|y, y\| \\ &= 2\|x, y\| + 2\|y, x\| \end{aligned}$$

Sostituendo y con iy otteniamo

$$\|x + iy\|^2 - \|x - iy\|^2 = 2\|x, iy\| + 2\|iy, x\| = -2i\|x, y\| + 2i\|y, x\|$$

per cui

$$\begin{aligned} \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 &= 2\|x, y\| + 2\|y, x\| \\ &\quad + 2\|x, y\| - 2\|y, x\| \\ &= 4\|x, y\| \end{aligned}$$

Lemma 4.11. *W sia uno spazio vettoriale normato su \mathbb{K} . Allora l'applicazione $\|\cdot\| : X \rightarrow \mathbb{R}$ è uniformemente continua.*

Dimostrazione. Ciò segue dalla disuguaglianza

$$|d(x, y) - d(u, v)| \leq d(x, u) + d(y, v)$$

valida in ogni spazio metrico.

Lemma 4.12. *W sia uno spazio vettoriale normato su \mathbb{K} in cui vale la legge del parallelogramma. Per $x, y \in W$ poniamo*

$\varphi(x, y) := \|x + y\|^2 - \|x - y\|^2$. Allora per ogni $x, y, z \in W$ valgono le seguenti relazioni:

- (1) $\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$.
- (2) $\varphi(x, y) = \varphi(y, x)$.
- (3) $\varphi(ix, y) = -\varphi(x, iy)$
 $\varphi(ix, iy) = \varphi(x, y)$
 nel caso $\mathbb{K} = \mathbb{C}$.

Dimostrazione. Abbiamo

$$(1) \|x + y + z\|^2 = 2\|x + z\|^2 + 2\|y\|^2 - \|x + z - y\|^2$$

Però anche

$$\|x + y + z\|^2 = 2\|y + z\|^2 + 2\|x\|^2 - \|y + z - x\|^2$$

per cui

$$\|x + y + z\|^2 = \|x + z\|^2 + \|y\|^2 + \|y + z\|^2 + \|x\|^2 - \frac{1}{2}\|x + z - y\|^2 - \frac{1}{2}\|y + z - x\|^2$$

Sostituendo z con $-z$ otteniamo

$$\|x + y - z\|^2 = \|x - z\|^2 + \|y\|^2 + \|y - z\|^2 + \|x\|^2 - \frac{1}{2}\|x - z - y\|^2 - \frac{1}{2}\|y - z - x\|^2$$

Sottraendo le due relazioni e tenendo conto che $\| -v \| = \|v\|$ per ogni $v \in W$ si ha

$$\|x + y + z\|^2 - \|x + y - z\|^2 = \|x + z\|^2 + \|y + z\|^2 - \|x - z\|^2 - \|y - z\|^2$$

ovvero

$$\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$$

(2) Chiaro.

(3) Siccome $|i| = 1$ abbiamo

$$\begin{aligned} \varphi(ix, y) &= \|ix + y\|^2 - \|ix - y\|^2 = \|i(x - iy)\|^2 + \|i(x + iy)\|^2 \\ &= \|x - iy\|^2 - \|x + iy\|^2 = -\varphi(x, iy) \end{aligned}$$

$$\varphi(ix, iy) = \|ix + iy\|^2 - \|ix - iy\|^2 = \|x + y\|^2 - \|x - y\|^2 = \varphi(x, y)$$

Proposizione 4.13. *W sia uno spazio vettoriale normato su \mathbb{K} in cui vale la legge del parallelogramma. Se definiamo*

$$\|x, y\| := \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2)$$

nel caso $\mathbb{K} = \mathbb{R}$ e

$$\|x, y\| := \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2)$$

nel caso $\mathbb{K} = \mathbb{C}$, allora otteniamo uno spazio vettoriale con prodotto scalare in cui $\|x\|^2 = \|x, x\|$ per ogni $x \in W$.

Dimostrazione. Dimostriamo le cinque proprietà dell'oss. 4.2. Siano $x, y, z \in V$. Ponendo $\varphi(x, y) := \|x + y\|^2 - \|x - y\|^2$ come nel lemma 4.12, abbiamo per definizione

$$4\|x, y\| = \varphi(x, y) + i\varphi(x, iy) \text{ nel caso } \mathbb{K} = \mathbb{C}$$

$$4\|x, y\| = \varphi(x, y) \text{ nel caso } \mathbb{K} = \mathbb{R}$$

(1) Dal lemma 4.12 segue che $\|x + y, z\| = \|x, z\| + \|y, z\|$.

(2) Sia $\lambda \in \mathbb{Z}$. Allora per il punto (1) si ha che $\|\lambda x, y\| = \lambda\|x, y\|$.

Se adesso $\lambda = \frac{m}{n}$ con $m, n \in \mathbb{Z}$ e $n \neq 0$, allora

$$\left\| \frac{m}{n}x, y \right\| = m \left\| \frac{1}{n}x, y \right\| = \frac{m}{n}n \left\| \frac{1}{n}x, y \right\| = \frac{m}{n} \left\| n \frac{1}{n}x, y \right\| = \frac{m}{n} \|x, y\|$$

per cui $\|\lambda x, y\| = \lambda \|x, y\|$ per ogni $\lambda \in \mathbb{Q}$.

Dalla continuità della norma (lemma 4.11) segue che

$$\|\lambda x, y\| = \lambda \|x, y\| \text{ per ogni } \lambda \in \mathbb{R}.$$

Nel caso $\mathbb{K} = \mathbb{C}$ abbiamo inoltre

$$\begin{aligned} 4\|ix, y\| &= \varphi(ix, y) + i\varphi(ix, iy) = -\varphi(x, iy) + i\varphi(x, y) \\ &= i(\varphi(x, y) + i\varphi(x, iy)) = 4i\|x, y\| \end{aligned}$$

e usando anche il punto (1), segue immediatamente che

$$\|\lambda x, y\| = \lambda \|x, y\| \text{ per ogni } \lambda \in \mathbb{C}.$$

(3) Nel caso $\mathbb{K} = \mathbb{R}$ abbiamo

$$4\|x, y\| = \varphi(x, y) = \varphi(y, x) = \|y, x\|$$

Nel caso $\mathbb{K} = \mathbb{C}$ invece, usando ancora il lemma 4.12,

$$\begin{aligned} 4\|y, x\| &= \varphi(y, x) + i\varphi(y, ix) = \varphi(x, y) + i\varphi(ix, y) \\ &= \varphi(x, y) - i\varphi(x, iy) = 4\overline{\|x, y\|} \end{aligned}$$

(4) Dal punto (3) segue in particolare che $\|x, x\|$ è reale anche nel caso $\mathbb{K} = \mathbb{C}$. Perciò

$$4\|x, x\| = \|x + x\|^2 - \|x - x\|^2 = 4\|x\|^2 \geq 0$$

(5) Da ciò segue inoltre che $\|x, x\| = 0$ implica $x = 0$ e che $\|x, x\|^2 = \|x\|^2$.

Proposizione 4.14. *L'applicazione $\| \cdot \| : V \times V \rightarrow \mathbb{K}$ è continua, ma non uniformemente continua (tranne nel caso banale $V = 0$).*

Dimostrazione. (1) La continuità segue dalla disuguaglianza

$$\begin{aligned} \left| \|u, v\| - \|x, y\| \right| &= \left| \|u - x, v\| + \|x, v\| - \|x, y\| \right| \\ &= \left| \|u - x, v\| + \|x, v - y\| \right| \leq \left| \|u - x, v\| \right| + \left| \|x, v - y\| \right| \\ &\leq \|u - x\| \|v\| + \|x\| \|v - y\| \end{aligned}$$

(2) Il prodotto scalare invece non è uniformemente continuo per la stessa ragione per cui non è uniformemente continua l'applicazione $\bigcirc_x x^2 : \mathbb{R} \rightarrow \mathbb{R}$. Altrimenti per ogni $\varepsilon > 0$ dovrebbe esistere un $\delta > 0$ tale che per $x, y, u, v \in V$ si abbia sempre

$$\|u - x\| \leq \delta \text{ e } \|v - y\| \leq \delta \Rightarrow \left| \|u, v\| - \|x, y\| \right| < \varepsilon$$

Però se scegliamo $w \in V$ con $\|w\| = 1$ (un tale w esiste sempre se $V \neq 0$), $x = tw$ con $t > 0$, $y = 0$, $u = x + \delta w$, $v = \delta w$, allora

$$\|u - x\| = \|\delta w\| = \delta, \quad \|v - y\| = \|\delta w\| = \delta,$$

mentre

$$\begin{aligned} \left| \|u, v\| - \|x, y\| \right| &= \left| \|u, v\| \right| = \left| \|x + \delta w, \delta w\| \right| = \left| \|(t + \delta)w, \delta w\| \right| \\ &= \delta(t + \delta)\|w\| = \delta(t + \delta) \end{aligned}$$

e ciò tende a ∞ per $t \rightarrow \infty$.

Definizione 4.15. V si chiama uno *spazio di Hilbert*, se V è completo nella metrica indotta dalla norma (cfr. oss. 4.8).

Nota 4.16. (1) Siano $a, b \in \mathbb{R}$ con $a < b$ e $V := C([a, b], \mathbb{K})$ lo spazio delle funzioni continue $[a, b] \rightarrow \mathbb{K}$ con il prodotto scalare

$$\|f, g\| := \int_a^b f(t) \overline{g(t)} dt$$

Allora V è uno spazio vettoriale con prodotto scalare, ma non è uno spazio di Hilbert.

(2) Sia V lo spazio di tutte le successioni $x = (x_0, x_1, x_2, \dots) \in \mathbb{K}^{\mathbb{N}}$ per le quali $x_n = 0$ per $n \gg 0$. Se poniamo $\|x, y\| := \sum_{n=0}^{\infty} x_n \overline{y_n}$, allora V diventa uno spazio vettoriale con prodotto scalare che non è uno spazio di Hilbert.

Dimostrazione. Testi di analisi funzionale.

Definizione 4.17. Due elementi $x, y \in V$ si dicono *ortogonali*, se $\|x, y\| = 0$. In tal caso scriviamo anche $x \perp y$.

Più in generale due insiemi $A, B \subset V$ si dicono *ortogonali* se $a \perp b$ per ogni $a \in A$ ed ogni $b \in B$. In tal caso scriveremo $A \perp B$.

Proposizione 4.18. Siano $x_1, \dots, x_m \in V$ e $x_i \perp x_j$ per $i \neq j$. Allora

$$\|x_1 + \dots + x_m\|^2 = \|x_1\|^2 + \dots + \|x_m\|^2$$

Dimostrazione. Scriviamo \sum_k invece di $\sum_{k=1}^m$. Si ha che

$$\begin{aligned} \left\| \sum_k x_k \right\|^2 &= \left\| \sum_k x_k, \sum_j x_j \right\| = \sum_k \sum_j \|x_k, x_j\| \\ &= \sum_k \|x_k, x_k\| = \sum_k \|x_k\|^2 \end{aligned}$$

Proposizione 4.19. Sia $x = (x_0, x_1, x_2, \dots) \in V^{\mathbb{N}}$ una successione di elementi di V tale che siano soddisfatte le seguenti condizioni:

(1) $x_i \perp x_j$ per $i \neq j$.

(2) $\sum_{k=0}^{\infty} x_k$ converge.

$$\text{Allora } \left\| \sum_{k=0}^{\infty} x_k \right\|^2 = \sum_{k=0}^{\infty} \|x_k\|^2.$$

Dimostrazione. Per $m \in \mathbb{N}$ siano $s_m := \sum_{k=0}^m x_k$ e $t_m := \sum_{k=0}^m \|x_k\|^2$.

Per ipotesi esiste $v := \sum_{k=0}^{\infty} x_k = \lim_{m \rightarrow \infty} s_m$. Dalla prop. 4.18 e dal lemma 4.11 segue che

$$\bigcirc_m t_m = \bigcirc_m \sum_{k=0}^m \|x_k\|^2 = \bigcirc_m \left\| \sum_{k=0}^{\infty} x_k \right\|^2 = \bigcirc_m \|s_m\|^2 \longrightarrow \|v\|^2$$

Ciò significa $\sum_{k=0}^{\infty} \|x_k\|^2 = \left\| \sum_{k=0}^{\infty} x_k \right\|^2$.

Definizione 4.20. Un sottoinsieme $A \subset V$ si dice *ortogonale*, se $A \neq \emptyset$ ed $a \perp b$ per ogni $a, b \in A$ con $a \neq b$.

Se in più $\|a\| = 1$ per ogni $a \in A$, allora A si dice *ortonormale*.

Osservazione 4.21. A sia un sottoinsieme ortogonale di V con $0 \notin A$. Allora A è linearmente indipendente.

Dimostrazione. Siano a_1, \dots, a_m elementi distinti di A e $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ tali che $\lambda_1 a_1 + \dots + \lambda_m a_m = 0$. Per ogni j allora

$$0 = \|\lambda_1 a_1 + \dots + \lambda_m a_m, a_j\| = \|\lambda_j a_j, a_j\| = \lambda_j \|a_j, a_j\|$$

e ciò implica $\lambda_j = 0$ perché per ipotesi $\|a_j, a_j\| \neq 0$.

5. Approssimazione in uno spazio di Banach

Situazione 5.1. V sia uno spazio vettoriale normato su \mathbb{K} .

Seguiamo Hirzebruch/Scharlau, pagg.72-80, e Schröder, pagg. 19-20.

Osservazione 5.2. Siano $x, y, u, v \in V$. Allora

$$\| \|x + y\| - \|u + v\| \| \leq \|x - u\| + \|y - v\|.$$

Dimostrazione. Applicando la disuguaglianza usata nel lemma 4.11 abbiamo

$$|d(x, -y) - d(u, -v)| \leq d(x, u) + d(-y, -v)$$

ovvero

$$\| \|x + y\| - \|u + v\| \| \leq \|x + u\| + \|y - v\|$$

Definizione 5.3. V si dice *uniformemente convesso*, se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che per $x, y \in (\|V\| \leq 1)$ valga l'implicazione

$$\|x - y\| \geq \varepsilon \implies \|x + y\| \leq 2 - \delta$$

Lemma 5.4. *Le seguenti condizioni sono equivalenti:*

- (1) V è uniformemente convesso.
- (2) Per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per $x, y \in (\|V\| = 1)$ valga l'implicazione $\|x - y\| \geq \varepsilon \implies \|x + y\| \leq 2 - \delta$.
- (3) Per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per $x, y \in (\|V\| = 1)$ valga l'implicazione $\|x + y\| > 2 - \delta \implies \|x - y\| < \varepsilon$.
- (4) Se $\bigcirc_n x_n$ e $\bigcirc_n y_n$ sono successioni in $(\|V\| = 1)$ tali che $\bigcirc_n \|x_n + y_n\| \longrightarrow 2$, allora $\bigcirc_n (x_n - y_n) \longrightarrow 0$.
- (5) Se $\bigcirc_n x_n$ e $\bigcirc_n y_n$ sono successioni in V tali che $\overline{\lim}_{n \rightarrow \infty} \|x_n\| \leq 1, \overline{\lim}_{n \rightarrow \infty} \|y_n\| \leq 1$ e $\bigcirc_n \|x_n + y_n\| \longrightarrow 2$, allora $\bigcirc_n (x_n - y_n) \longrightarrow 0$.

Dimostrazione. (1) \implies (2) \implies (3) \implies (4) Chiaro.

(4) \implies (5): (A) Dimostriamo prima che $\bigcirc_n \|x_n\| \longrightarrow 1$. Se non fosse così, dalle ipotesi seguirebbe che esiste una sottosuccessione $\bigcirc_k x_{n_k}$ tale che $\bigcirc_k x_{n_k} \longrightarrow \alpha$ con $0 \leq \alpha < 1$. Ma allora si avrebbe

$$\begin{aligned} 2 &= \lim_{k \rightarrow \infty} \|x_{n_k} + y_{n_k}\| = \overline{\lim}_{k \rightarrow \infty} \|x_{n_k} + y_{n_k}\| \leq \overline{\lim}_{k \rightarrow \infty} \|x_{n_k}\| + \overline{\lim}_{k \rightarrow \infty} \|y_{n_k}\| \\ &= \alpha + \overline{\lim}_{k \rightarrow \infty} \|y_{n_k}\| \leq \alpha + 1 < 2 \end{aligned}$$

ma contraddizione. Per la stessa ragione $\bigcirc_n \|y_n\| \longrightarrow 1$.

(B) In particolare $x_n, y_n \neq 0$ per $n \gg 0$ e possiamo assumere che $x_n, y_n \neq 0$ per ogni n . Siano $x'_n := \frac{x_n}{\|x_n\|}$ ed $y'_n := \frac{y_n}{\|y_n\|}$. Dal punto (A)

segue che $\bigcirc_n \|x'_n - x_n\| \rightarrow 0, \bigcirc_n \|y'_n - y_n\| \rightarrow 0$. Dall'oss. 5.2 abbiamo

$$\| \|x'_n + y'_n\| - \|x_n + y_n\| \| \leq \|x'_n - x_n\| + \|y'_n - y_n\|$$

e ciò tende a 0, come abbiamo appena osservato.

Perciò $\lim_{n \rightarrow \infty} \|x'_n + y'_n\| = \lim_{n \rightarrow \infty} \|x_n + y_n\| = 2$ e per ipotesi, essendo x'_n ed y'_n elementi di ($\|V\| = 1$), ciò implica $\bigcirc_n (x'_n - y'_n) \rightarrow 0$.

Dalla decomposizione

$$x_n - y_n = (x_n - x'_n) + (x'_n - y'_n) + (y'_n - y_n)$$

segue adesso $\bigcirc_n (x_n - y_n) \rightarrow 0$.

(5) \implies (1): Assumiamo che esista un $\varepsilon > 0$ tale che per ogni $n \in \mathbb{N}+1$ esistano $x_n, y_n \in (\|V\| \leq 1)$ con $\|x_n - y_n\| \geq \varepsilon$ ed $\|x_n + y_n\| > 2 - 1/n$. Siccome $\|x_n + y_n\| \leq \|x_n\| + \|y_n\| \leq 2$, ciò implica $\bigcirc_n \|x_n + y_n\| \rightarrow 2$.

Per ipotesi quindi $\bigcirc_n (x_n - y_n) \rightarrow 0$, ma ciò non è possibile essendo $\|x_n - y_n\| \geq \varepsilon$ per ogni n .

Lemma 5.5. *V sia uniformemente convesso ed $x, y \in (\|V\| \leq 1)$ tali che $\|x + y\| \geq 2$. Allora $x = y$.*

Dimostrazione. Sia infatti $x \neq y$. Allora $\|x - y\| > 0$ e dalla def. 5.3 abbiamo che $\|x + y\| < 2$.

Definizione 5.6. *V si dice strettamente normato, se per $x, y \in V \setminus 0$ con $\|x + y\| = \|x\| + \|y\|$ esiste $\lambda > 0$ tale che $y = \lambda x$.*

Si noti che allora $\|y\| = \lambda\|x\|$ e quindi necessariamente $\lambda = \frac{\|y\|}{\|x\|}$.

Proposizione 5.7. *V sia uniformemente convesso. Allora V è strettamente normato.*

Dimostrazione. Siano $x, y \in V \setminus 0$ tali che $\|x + y\| = \|x\| + \|y\|$.

Sia ad esempio $0 < \|x\| \leq \|y\|$ e quindi

$$\left| \frac{1}{\|x\|} - \frac{1}{\|y\|} \right| = \left| \frac{\|y\| - \|x\|}{\|x\|\|y\|} \right| = \frac{\|y\| - \|x\|}{\|x\|\|y\|}$$

Allora $\left\| \frac{x}{\|x\|} + \frac{y}{\|y\|} \right\| \leq \left\| \frac{x}{\|x\|} + \frac{y}{\|y\|} \right\| + \left\| \frac{y}{\|x\|} - \frac{y}{\|y\|} \right\|$ e quindi

$$\begin{aligned} \left\| \frac{x}{\|x\|} + \frac{y}{\|y\|} \right\| &\geq \left\| \frac{x}{\|x\|} + \frac{y}{\|x\|} \right\| - \left\| \frac{y}{\|x\|} - \frac{y}{\|y\|} \right\| \\ &= \left\| \frac{x + y}{\|x\|} \right\| - \left\| \frac{\|y\|y - \|x\|y}{\|x\|\|y\|} \right\| \\ &= \frac{\|x\| + \|y\|}{\|x\|} - \frac{\|y\| - \|x\|}{\|x\|\|y\|} \|y\| \\ &= \frac{\|x\| + \|y\| - \|y\| + \|x\|}{\|x\|} = \frac{2\|x\|}{\|x\|} = 2 \end{aligned}$$

Dal lemma 5.5 segue che $\frac{x}{\|x\|} = \frac{y}{\|y\|}$.

Proposizione 5.8. *Uno spazio con prodotto scalare è uniformemente convesso e quindi strettamente normato.*

Dimostrazione. V sia uno spazio con prodotto scalare. Per la prop. 4.9 in V vale la legge del parallelogramma.

Per $x, y \in (\|V\| = 1)$ si ha quindi $\|x + y\|^2 = 4 - \|x - y\|^2$ e quindi, se $\varepsilon > 0$ e $\|x - y\| \geq \varepsilon$, allora $\|x + y\| \leq \sqrt{4 - \varepsilon^2}$.

Per $\varepsilon_0 := \min(\varepsilon, 1/2)$ abbiamo allora $\delta := 2 - \sqrt{4 - \varepsilon_0^2} > 0$ e

$$\|x + y\| \leq \sqrt{4 - \varepsilon^2} \leq \sqrt{4 - \varepsilon_0^2} = 2 - \delta$$

Corollario 5.9. \mathbb{K}^n è uniformemente convesso e strettamente normato rispetto alla norma euclidea.

Definizione 5.10. (X, d) sia uno spazio metrico. Per $x \in X$ e $A \subset X$ definiamo $d(x, A) := \inf \{d(x, a) \mid a \in A\}$.

Osservazione 5.11. Siano (X, d) uno spazio metrico, $x \in X$ ed $A \subset X$. Allora $x \in \bar{A} \iff d(x, A) = 0$.

Dimostrazione. Ad es. Engelking, pag. 254.

Osservazione 5.12. Siano (X, d) uno spazio metrico ed A un sottoinsieme completo di X . Allora A è chiuso in X .

Dimostrazione. Ad es. Engelking, pag. 270.

Lemma 5.13. V sia strettamente normato e C un sottoinsieme convesso di V . Siano $x \in V$ ed $a, b \in C$ tali che $\|x - a\| = \|x - b\| = d(x, C)$.

Allora $a = b$.

Dimostrazione. Sia $\rho := d(x, C)$. Per la convessità di C abbiamo $\frac{a+b}{2} \in C$ e inoltre

$$\left\|x - \frac{a+b}{2}\right\| = \frac{1}{2}\|x - a + x - b\| \leq \frac{1}{2}(\|x - a\| + \|x - b\|) = \rho$$

cosicché la definizione di $d(x, C)$ implica

$$\|x - a + x - b\| = 2\rho = \|x - a\| + \|x - b\|$$

Ma V è strettamente normato e quindi, se $x \neq a, b$, abbiamo

$$x - b = \frac{\|x - b\|}{\|x - a\|} (x - a) = x - a$$

e quindi $a = b$.

Il caso $x = a$ o $x = b$ è banale, perché allora $x = a = b$.

Teorema 5.14. V sia uniformemente convesso e C un sottoinsieme convesso, completo e non vuoto di V . Allora per ogni $x \in V$ esiste esattamente un punto $p \in C$ tale che $d(x, C) = \|x - p\|$.

Dimostrazione. (1) L'unicità segue dal lemma 5.13

(2) Siccome possiamo sostituire V con $V - x$, possiamo assumere che $x = 0$.

(3) Sia $\rho := d(0, C)$

(4) Sia $\rho = 0$. Per l'osservazione 5.11 allora $0 \in \overline{C} \stackrel{5.12}{=} C$ e possiamo porre $p = 0$.

(5) Sia perciò $\rho > 0$. Allora possiamo sostituire C con C/ρ e supporre quindi che $\rho = 1$. Per la definizione di $d(0, C)$ esiste una successione $\bigcirc_n c_n$ di elementi di C tale che $\lim_{n \rightarrow \infty} \|c_n\| = 1$. Per ogni $n, m \in \mathbb{N}$ allora $\frac{c_n + c_m}{2} \in C$ (ciò segue dalla convessità di C) e quindi

$$1 \leq \left\| \frac{c_n + c_m}{2} \right\| \leq \frac{\|c_n\|}{2} + \frac{\|c_m\|}{2} \quad (*)$$

Dimostriamo che $\bigcirc_n c_n$ è una successione di Cauchy. Altrimenti esisterebbero un $\varepsilon > 0$ e due sottosuccessioni $\bigcirc_k c_{n_k}$ e $\bigcirc_k c_{m_k}$ tali che $\|c_{n_k} - c_{m_k}\| \geq \varepsilon$ per ogni k .

Dalla relazione (*) e dall'ipotesi $\bigcirc_n \|c_n\| \rightarrow 1$ segue però che $\bigcirc_k \|c_{n_k} + c_{m_k}\| \rightarrow 2$. Ma ciò è in contraddizione con la condizione (5) del lemma 5.4.

(6) Per la completezza di C esiste quindi un punto $p \in C$ tale che $\bigcirc_n c_n \rightarrow p$. Dal lemma 4.11 segue che $\rho = \lim_{n \rightarrow \infty} \|x - c_n\| = \|x - p\|$.

Corollario 5.15. *V sia uno spazio di Banach uniformemente convesso e C un sottoinsieme convesso, chiuso e non vuoto di V . Allora per ogni $x \in V$ esiste esattamente un punto $p \in C$ tale che $d(x, C) = \|x - p\|$.*

Corollario 5.16. *V sia uno spazio di Hilbert e C un sottoinsieme convesso, chiuso e non vuoto di V . Allora per ogni $x \in V$ esiste esattamente un punto $p \in C$ tale che $d(x, C) = \|x - p\|$.*

Dimostrazione. Per la prop. 5.8 V è uniformemente convesso. L'enunciato segue quindi dal corollario 5.15.

Osservazione 5.17. Su $V := \mathbb{R}_2$ introduciamo la norma $\|\cdot\|_1$ definita da $\|(x_1, x_2)\|_1 := |x_1| + |x_2|$.

Si dimostra facilmente che $l^1(2) := (V, \|\cdot\|_1)$ è uno spazio di Banach la cui topologia coincide con quella euclidea. Questo spazio però non è strettamente normato (e quindi nemmeno uniformemente convesso) e infatti non vale il teorema di approssimazione:

(1) Siano $x := (1, 0)$ e $y := (0, 1)$. Allora

$$\|x + y\|_1 = \|(1, 1)\|_1 = 2 = \|x\|_1 + \|y\|_1$$

ma $y \neq \lambda y$ per ogni $\lambda \in \mathbb{R}$. Perciò $l^1(2)$ non è strettamente normato.

(2) Consideriamo il segmento di retta

$$C := \{(x_1, x_2) \mid x_1, x_2 \in [0, 1] \text{ e } x_1 + x_2 = 1\}$$

Per ogni punto $p = (x_1, x_2) \in C$ allora $\|p\|_1 = |x_1| + |x_2| = x_1 + x_2 = 1$,
perciò $d(0, p)$ non dipende da $p \in C$.

6. Operatori lineari continui

Situazione 6.1. V e W siano spazi vettoriali normati.

Definizione 6.2. Per un'applicazione lineare $\varphi : V \rightarrow W$ sia

$$\|\varphi\| := \sup \{ \|\varphi v\| \mid v \in (\|v\| = 1) \} = \sup \left\{ \frac{\|\varphi v\|}{\|v\|} \mid v \in V \setminus 0 \right\}$$

Allora $\|\varphi\| \in [0, \infty]$.

φ si dice *limitata*, se $\|\varphi\| < \infty$.

Proposizione 6.3. Per un'applicazione lineare $\varphi : V \rightarrow W$ sono equivalenti:

(1) φ è uniformemente continua.

(2) φ è continua.

(3) φ è continua in 0.

(4) φ è limitata.

(5) Esiste un numero reale $A \geq 0$ tale che $\|\varphi v\| \leq A\|v\|$ per ogni $v \in V$.

Dimostrazione. (1) \implies (2) \implies (3): Chiaro.

(3) \implies (4): L'ipotesi implica che esiste un $\delta > 0$ tale che per $\|v\| \leq \delta$ si ha $\|\varphi v\| \leq 1$.

Sia $\|v\| = 1$. Allora $\|\delta v\| = \delta$ e quindi $\delta\|\varphi v\| = \|\varphi\delta v\| \leq 1$,

cosicché $\|\varphi v\| \leq 1/\delta$. Perciò $\|\varphi\| \leq 1/\delta$.

(4) \implies (5): Siano $A := \|\varphi\| < \infty$ e $v \in V$. La disuguaglianza $\|\varphi v\| \leq A\|v\|$ è banalmente soddisfatta per $v = 0$. Per $v \neq 0$ abbiamo

$$\frac{\|\varphi v\|}{\|v\|} \leq A \text{ e quindi } \|\varphi v\| \leq A\|v\|.$$

(5) \implies (1): Siano $u, v \in V$. Allora $\|\varphi u - \varphi v\| = \|\varphi(u - v)\| \leq A\|u - v\|$

e vediamo che φ è uniformemente continua.

Definizione 6.4. $\mathcal{L}(V, W)$ sia l'insieme delle applicazioni lineari continue $V \rightarrow W$. Poniamo inoltre $\mathcal{L}(V) := \mathcal{L}(V, V)$.

È chiaro che $\mathcal{L}(V, W)$ è uno spazio vettoriale.

Lo spazio $V' := \mathcal{L}(V, \mathbb{K})$ si chiama *spazio duale* di V .

Proposizione 6.5. Se per $\varphi \in \mathcal{L}(V, W)$ definiamo $\|\varphi\|$ come nella def. 6.2, allora $\mathcal{L}(V, W)$ denota uno spazio vettoriale normato.

Se W è uno spazio di Banach, anche $\mathcal{L}(V, W)$ è uno spazio di Banach.

Dimostrazione. Facile verifica, ad es. Hirzebruch/Scharlau, pagg. 34-35.

7. Spazi di Hilbert

Situazione 7.1. $(V, \|\cdot\|)$ sia uno spazio di Hilbert su \mathbb{K} .

Definizione 7.2. X sia un sottoinsieme di V . Poniamo

$$X^\perp := \{v \in V \mid v \perp x \text{ per ogni } x \in X\}.$$

Proposizione 7.3. X sia un sottoinsieme di V . Allora X^\perp è un sottospazio vettoriale chiuso di V .

Dimostrazione. Ciò segue dalla linearità e dalla continuità del prodotto scalare.

Osservazione 7.4. $V^\perp = 0$.

Dimostrazione. Sia $v \in V^\perp$. Allora $\|v, v\| = 0$ e quindi $v = 0$.

Osservazione 7.5. X ed Y siano sottoinsiemi di V tali che $X \perp Y$. Allora $X \cap Y = \emptyset$ oppure $X \cap Y = 0$.

Se X e Y sono sottospazi vettoriali, si ha quindi $X \cap Y = 0$.

Dimostrazione. Sia $v \in X \cap Y$. Per ipotesi allora $\|v, v\| = 0$ e ciò implica $v = 0$.

Corollario 7.6. W sia un sottospazio vettoriale di V . Allora $W \cap W^\perp = 0$.

Definizione 7.7. I sia un insieme e W_α un sottospazio vettoriale di V per ogni $\alpha \in I$ (diciamo allora che $\bigcirc_{\alpha \in I} W_\alpha$ è una famiglia di sottospazi vettoriali di V). Diciamo che V è somma ortogonale di W_α e scriviamo $V = \boxplus_{\alpha \in I} W_\alpha$, se sono soddisfatte le seguenti condizioni:

$$(1) V = \sum_{\alpha \in I} W_\alpha := \text{SV} \left(\bigcup_{\alpha \in I} W_\alpha \right).$$

$$(2) W_\alpha \perp W_\beta \text{ per } \alpha \neq \beta.$$

Nel caso $I = \{1, \dots, k\}$ scriviamo anche $V = W_1 \boxplus \dots \boxplus W_k$.

Osservazione 7.8. $\bigcirc_{\alpha} W_\alpha$ sia una famiglia di sottospazi vettoriali di V tale che $V = \boxplus_{\alpha} W_\alpha$. Allora $V = \bigoplus_{\alpha} W_\alpha$, cioè $W_\alpha \cap W_\beta = 0$ per $\alpha \neq \beta$.

Dimostrazione. Ciò segue dall'oss. 7.5.

Osservazione 7.9. W_1 e W_2 siano sottospazi vettoriali di V con $V = W_1 \boxplus W_2$. Allora $W_2 = W_1^\perp$.

Dimostrazione. (1) È chiaro che $W_2 \subset W_1^\perp$.

(2) Sia $v \in W_1^\perp$. Per ipotesi possiamo scrivere v della forma $v = w_1 + w_2$ con $w_1 \in W_1$ e $w_2 \in W_2$. Allora

$$0 = \|v, w_1\| = \|w_1 + w_2, w_1\| = \|w_1, w_1\| + \|w_2, w_1\| = \|w_1, w_1\|$$

per cui $w_1 = 0$ e quindi $v = w_2 \in W_2$.

Ciò mostra che $W_1^\perp \subset W_2$.

Corollario 7.10. $\bigcirc_{\alpha \in I} W_\alpha$ sia una famiglia di sottospazi vettoriali di V tali che $V = \boxplus_{\alpha \in I} W_\alpha$. Per ogni $\beta \in I$ allora $W_\beta^\perp = \boxplus_{\alpha \in I \setminus \beta} W_\alpha$.

Definizione 7.11. W sia un sottospazio vettoriale chiuso di V . Per ogni $x \in V$ esiste allora esattamente un punto $x_W \in W$ tale che $d(x, W) = \|x - x_W\|$, come sappiamo dal cor. 5.16.

Otteniamo così un'applicazione $\pi_W := \bigcirc_x x_W : V \rightarrow V$ che si chiama la *proiezione ortogonale* su W .

Si noti che come codominio di π_W abbiamo scelto V anche se ovviamente l'immagine di π_W coincide con W .

Lemma 7.12. W sia un sottospazio vettoriale chiuso di V . Per $x, p \in V$ sono allora equivalenti:

- (1) $p = \pi_W(x)$.
- (2) $p \in W$ e $x - p \in W^\perp$.

Dimostrazione. (1) \implies (2): Sia $p := \pi_W(x)$. Allora $p \in W$. Sia $w \in W$.

Se $w = 0$, chiaramente $\|x - p, w\| = 0$.

Sia perciò $w \neq 0$. Per $\lambda := \frac{\|x - p, w\|}{\|w\|^2}$ si ha allora $p + \lambda w \in W$ e quindi

$$\begin{aligned} \|x - p\|^2 &\leq \|x - p - \lambda w\|^2 \\ &= \|x - p\|^2 - 2\lambda \|x - p, w\| + \lambda^2 \|w\|^2 \\ &= \|x - p\|^2 - \frac{\|w, x - p\| \|x - p, w\|}{\|w\|^2} - \frac{\|x - p, w\| \|w, x - p\|}{\|w\|^2} \\ &\quad + \frac{\|x - p, w\| \|w, x - p\|}{\|w\|^2} \\ &= \|x - p\|^2 - \frac{\| \|x - p, w\| \|^2}{\|w\|^2} \end{aligned}$$

per cui $\|x - p, w\| = 0$ e quindi $x - p \perp w$.

(2) \implies (1): Sia $p \in W$ e $x - p \perp W$. Per $w \in W$ allora $p - w \in W$ e quindi $x - p \perp p - w$. Dalla prop. 4.18 segue allora

$$\begin{aligned} \|x - w\|^2 &= \|x - p + p - w\|^2 \\ &= \|x - p\|^2 + \|p - w\|^2 \geq \|x - p\|^2 \end{aligned}$$

Proposizione 7.13. W sia un sottospazio vettoriale chiuso di V . Allora l'applicazione $\pi_W : V \rightarrow V$ è lineare e continua e quindi un elemento di $\mathcal{L}(V)$.

Per $W \neq 0$ si ha $\|\pi_W\| = 1$.

Dimostrazione. (1) Siano $x, y \in V$ e $p := \pi_W(x), q := \pi_W(y)$.

Per $\alpha, \beta \in \mathbb{K}$ allora $\alpha p + \beta q \in W$ e

$$(\alpha x + \beta y) - (\alpha p + \beta q) = \alpha(x - p) + \beta(y - q) \in W^\perp + W^\perp = W^\perp$$

cosicché dal lemma 7.12 segue

$$\pi_W(\alpha x + \beta y) = \alpha p + \beta q = \alpha \pi_W(x) + \beta \pi_W(y)$$

Ciò mostra che π_W è un'applicazione lineare.

(2) Siano $x \in V$ e $p := \pi_W(x)$. Allora

$$\|x\|^2 = \|x - p + p\|^2 = \|x - p\|^2 + \|p\|^2 \geq \|p\|^2 \text{ e vediamo che } \|\pi_W\| \leq 1.$$

Per $W \neq 0$ scegliamo un $w \in W \setminus 0$ e siccome allora

$$\|\pi_W(w)\| = \|w\| \text{ vediamo che } \|\pi_W\| = 1.$$

Osservazione 7.14. Siano W un sottospazio vettoriale chiuso di V e $P := \pi_W$. Allora

(1) $Pw = w$ per ogni $w \in W$.

(2) $Px = 0$ per ogni $x \in W^\perp$.

(3) $P^2 = P$.

Dimostrazione. (1) Chiaro (cfr. 7.11).

(2) Ciò segue dal lemma 7.12. Infatti se $x \in W^\perp$, allora per $p = 0$ si ha $p \in W$ e $x - p \in W^\perp$.

(3) Segue dal punto (1).

Definizione 7.15. Sia $\varphi : V \rightarrow V$ un'applicazione lineare. φ si dice *simmetrica* (rispetto al prodotto scalare dato), se $\|\varphi x, y\| = \|x, \varphi y\|$ per ogni $x, y \in V$.

Per un noto teorema dell'analisi funzionale φ è allora automaticamente continua (teorema di Hellinger-Toeplitz, cfr. Heuser, pag. 244, oppure Schröder, pag 73).

Proposizione 7.16. Sia W un sottospazio vettoriale chiuso di V . Allora la proiezione π_W è simmetrica.

Dimostrazione. Siano $x, y \in V$ e $p := \pi_W(x), q := \pi_W(y)$. Dobbiamo dimostrare che $\|p, y\| = \|x, q\|$.

Dal lemma 7.12 abbiamo $\|x - p, q\| = \|p, y - q\| = 0$ e quindi

$$\|p, y\| = \|p, q\| = \|x, q\|$$

Corollario 7.17. Sia W un sottospazio vettoriale chiuso di V . Allora $V = W \oplus W^\perp$ e quindi anche $V = W \boxplus W^\perp$.

Dimostrazione. (1) Dal cor. 7.6 sappiamo che $W \cap W^\perp = 0$.

(2) Siano $x \in V$ e $p := \pi_W(x)$.

Per il lemma 7.12 allora $x - p \in W^\perp$ e quindi

$$x = p + (x - p) \in W + W^\perp$$

Proposizione 7.18. *Sia W un sottospazio vettoriale chiuso di V . Allora $W^{\perp\perp} = W$.*

Dimostrazione. Dal cor. 7.17 sappiamo che $V = W \boxplus W^\perp$, per cui per l'oss. 7.9 abbiamo $W = W^{\perp\perp}$.

Osservazione 7.19. W sia un sottospazio vettoriale di V per il quale esiste un sottospazio vettoriale U di V tale che $V = W \boxplus U$.

Allora W è chiuso.

Dimostrazione. Per il corollario 7.10 abbiamo che $W = U^\perp$. Per la prop. 7.3 W è chiuso.

Corollario 7.20. W_1 e W_2 siano sottospazi vettoriali chiusi di V . Allora $W_1 = W_2 \iff W_1^\perp = W_2^\perp$.

Dimostrazione. \implies : Chiaro.

\impliedby : Sia $W_1^\perp = W_2^\perp$. Dalla prop. 7.18 segue $W_1 = W_1^{\perp\perp} = W_2^{\perp\perp} = W_2$.

Osservazione 7.21. Sia X un sottoinsieme di V . Allora $\overline{X}^\perp = X^\perp$.

Dimostrazione. È chiaro che $\overline{X}^\perp \subset X^\perp$.

Siano $v \in X^\perp$ e $z \in \overline{X}$. Allora esiste una successione $\bigcirc_n x_n$ di elementi di X tale che $\bigcirc_n x_n \longrightarrow z$. Dalla continuità del prodotto scalare (prop. 6.3) abbiamo perciò $\bigcirc_n \|x_n, v\| \longrightarrow \|z, v\|$.

Ma per ipotesi $\|x_n, v\| = 0$ per ogni n e quindi anche $\|z, v\| = 0$.

Osservazione 7.22. W sia un sottospazio vettoriale di V . Allora anche \overline{W} è un sottospazio vettoriale di V .

Corollario 7.23. W sia un sottospazio vettoriale di V .

Allora $\overline{W} = V \iff W^\perp = 0$.

Dimostrazione. $\overline{W} = V \iff \overline{W}^\perp = V^\perp \stackrel{7.4}{\iff} \overline{W}^\perp = 0 \stackrel{7.21}{\iff} W^\perp = 0$.

Definizione 7.24. Un *proiettore* in V è un'applicazione lineare $P : V \longrightarrow V$ tale che $P^2 = P$.

Osservazione 7.25. $P : V \longrightarrow V$ sia un proiettore.

Allora $\text{im } P = \ker(\text{id} - P)$.

Dimostrazione. (1) Sia $x \in \text{im } P$, ad esempio $x = Pv$ con $v \in V$. Allora $(\text{id} - P)x = x - Px = Pv - P^2v = 0$.

(2) Sia $x \in \ker(\text{id} - P)$. Allora $x = Px \in \text{im } P$.

Corollario 7.26. $P : V \longrightarrow V$ sia un proiettore continuo. Allora $\text{im } P$ è un sottospazio vettoriale chiuso di V .

Dimostrazione. La continuità di P implica che $\ker(\text{id} - P)$ è chiuso. L'enunciato segue dall'oss. 7.25.

Osservazione 7.27. $P : V \longrightarrow V$ sia un proiettore simmetrico. Per ogni $x, z \in V$ allora $\|Px, Pz\| = \|x, Pz\|$.

Dimostrazione. La simmetria implica $\|Px, Pz\| = \|x, P^2z\|$ e ciò è uguale a $\|x, Pz\|$ perché per ipotesi $P^2 = P$.

Proposizione 7.28. $P : V \longrightarrow V$ sia un proiettore simmetrico. Allora $\text{im } P$ è un sottospazio vettoriale chiuso di V e $P = \pi_{\text{im } P}$.

Dimostrazione. Per il teorema di Hellinger-Toeplitz P è continua e quindi $W := \text{im } P$ è un sottospazio vettoriale chiuso di V per il cor. 7.26.

Per il lemma 7.12 è sufficiente dimostrare che $x - Px \in (\text{im } P)^\perp$ per ogni $x \in V$. Sia $y = Pv \in \text{im } P$. Allora

$$\begin{aligned} \|x - Px, y\| &= \|x - Px, Pv\| = \|x, Pv\| - \|Px, Pv\| \\ &\stackrel{7.27}{=} \|Px, Pv\| - \|Px, Pv\| = 0 \end{aligned}$$

Corollario 7.29. I seguenti insiemi coincidono:

- (1) $\{P : V \longrightarrow V \mid P \text{ è un proiettore simmetrico}\}$.
- (2) $\{\pi_W \mid W \text{ è un sottospazio vettoriale chiuso di } V\}$.

Osservazione 7.30. $P : V \longrightarrow V$ sia un proiettore. Allora:

- (1) Anche $(\text{id} - P)$ è un proiettore.
- (2) $\text{im } (\text{id} - P) = \ker P$.
- (3) Se P è simmetrico, anche $\text{id} - P$ è simmetrico.
- (4) $V = \text{im } P \oplus \text{im } (\text{id} - P) = \text{im } P \boxplus \ker P$.
- (5) Se P è simmetrico, allora $V = \text{im } P \boxplus \text{im } (\text{id} - P) = \text{im } P \boxplus \ker P$.

Dimostrazione. (1) $(\text{id} - P)^2 = \text{id} - 2P + P^2 = \text{id} - P$.

(2) Per il punto (1) possiamo applicare l'oss. 7.25, da cui segue $\text{im } (\text{id} - P) = \ker P$.

(3) P sia simmetrico ed $x, y \in V$. Allora

$$\begin{aligned} \|(\text{id} - P)x, y\| &= \|x - Px, y\| = \|x, y\| - \|Px, y\| \\ &\stackrel{7.16}{=} \|x, y\| - \|x, Py\| = \|x, y - Py\| = \|x, (\text{id} - P)y\| \end{aligned}$$

(4) Per ogni $x \in V$ si ha $x = Px + x - Px$, per cui è chiaro che $V = \text{im } P + \text{im } (\text{id} - P)$.

Dimostriamo che $\text{im } P \cap \text{im } (\text{id} - P) = 0$. Sia $x = Py = z - Pz$ con $y, z \in V$. Allora $x = Py = P^2y = Pz - P^2z = Pz - Pz = 0$.

(5) P sia simmetrico ed $x, y \in V$. Allora

$$\|Px, y - Py\| = \|Px, y\| - \|Px, Py\| = \|x, Py\| - \|x, Py\| = 0.$$

Corollario 7.31. W sia un sottospazio vettoriale chiuso di V .

Allora $W^\perp = \text{im}(\text{id} - \pi_W)$.

Dimostrazione. Dal cor. 7.17 sappiamo che $V = W \oplus W^\perp$ e dall'oss. 7.30 abbiamo che $V = \text{im} \pi_W \oplus \text{im}(\text{id} - \pi_W) = W \oplus \text{im}(\text{id} - \pi_W)$. Dall'oss. 7.9 segue che $\text{im}(\text{id} - \pi_W) = W^\perp$.

Proposizione 7.32. $P, Q : V \longrightarrow V$ siano proiettori simmetrici.

Se $\text{im} P = \text{im} Q$, allora $P = Q$.

Dimostrazione. Ciò segue dalla prop. 7.28.

Osservazione 7.33. Nel seguito in alcuni esempi identificheremo, per $A \in \mathbb{K}_n^m$, la matrice A con l'applicazione $\bigcirc_x Ax$.

Osservazione 7.34. La condizione che P e Q siano simmetrici nella prop. 7.32 è necessaria. Infatti per $V = \mathbb{K}^2$ e $P = \begin{pmatrix} 1 & \lambda \\ 0 & 0 \end{pmatrix}$ con $\lambda \in \mathbb{K}$ si ha $P^2 = P$.

Le immagini di tutti questi proiettori coincidono, ma solo per $\lambda = 0$ si tratta di una proiezione ortogonale.

Esempio 7.35. I proiettori simmetrici $P : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ con $\dim \text{im} P = 1$ sono dati esattamente dalle matrici della forma $\begin{pmatrix} \cos^2 t & \cos t \sin t \\ \cos t \sin t & \sin^2 t \end{pmatrix}$ con $t \in [0, 2\pi)$.

Dimostrazione. Sia W una retta passante per l'origine in \mathbb{R}^2 . Allora $W = \mathbb{R}e_1$ con $e_1 = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, $t \in \mathbb{R}$. La proiezione ortogonale su W è data da $P = \bigcirc_x \|x, e_1\| e_1$. In particolare abbiamo

$$P\delta_1 = \left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \right\| \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \cos t \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} \cos^2 t \\ \cos t \sin t \end{pmatrix}$$

$$P\delta_2 = \left\| \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \right\| \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \sin t \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} \cos t \sin t \\ \sin^2 t \end{pmatrix}$$

L'enunciato segue dalla prop. 7.32.

Proposizione 7.36. W sia un sottospazio vettoriale di V e $\dim V < \infty$.

Allora W è chiuso in V .

Dimostrazione. Scelta una base e_1, \dots, e_m di W , si dimostra che W è completo nello stesso modo in cui si dimostra che \mathbb{K}^n è completo. Per l'oss. 5.12 W è chiuso in V .

Proposizione 7.37. $P : V \longrightarrow V$ sia un proiettore. Allora sono equivalenti:

- (1) P è simmetrico.
- (2) $\|P\| = 1$.

(3) $\text{im } P \perp \ker P$.

Dimostrazione. (1) \implies (2): Per la prop. 7.28 $\text{im } P$ è un sottospazio vettoriale chiuso di V e per la prop. 7.13 abbiamo che $\|P\| = 1$.

(2) \implies (3): Sia $\|P\| = 1$.

(A) Dimostriamo prima che $(\ker P)^\perp \subset \text{im } P$.

Sia $x \in (\ker P)^\perp$. Siccome $x - Px \in \ker P$, abbiamo allora

$0 = \|x - Px, x\| = \|x\|^2 - \|Px, x\|$ e quindi $\|Px, x\| = \|x\|^2$. Vediamo in particolare che $\|Px, x\|$ è reale e ≥ 0 .

Usando l'ipotesi che $\|P\| = 1$ abbiamo inoltre

$$\|x\|^2 = \|Px, x\| \leq \|Px\| \|x\| \leq \|x\|^2$$

cosicché $\|x\| = \|Px\| = \sqrt{\|Px, x\|}$ (essendo ciò banalmente vero anche per $x = 0$). Ma allora, essendo $\|Px, x\| = \|x, Px\| = \|x, Px\|$,

$$\|x - Px\|^2 = \|x\|^2 - \|x, Px\| - \|Px, x\| + \|Px\|^2 = 0$$

e quindi $x = Px \in \text{im } P$.

(B) Dimostriamo che $\text{im } P \subset (\ker P)^\perp$.

Sia $y \in \text{im } P$. L'ipotesi $\|P\| = 1$ implica però che P è continua, per cui $\ker P$ è chiuso, cosicché $V = \ker P \oplus (\ker P)^\perp$ per il corollario 7.17. Perciò esistono $u \in \ker P$ e $v \in (\ker P)^\perp$ tali che $y = u + v$. Per ipotesi $y \in \text{im } P$, perciò $y = Pz$. Ciò implica $y = Py = Pu + Pv = Pv$. Per il punto (1) però $v \in \text{im } P$, per cui $Pv = v$, cosicché $y = v \in (\ker P)^\perp$.

(3) \implies (1): Assumendo che $\text{im } P \perp \ker P$ siano $x, y \in V$. Per l'oss. 7.30 possiamo scrivere $x = a + b, y = c + d$ con $a, c \in \text{im } P$ e $b, d \in \ker P$, avendo allora $Px = a, Py = c$ e quindi

$$\|Px, y\| = \|a, c + d\| = \|a, c\|$$

però anche

$$\|x, Py\| = \|a + b, c\| = \|a, c\|$$

8. Il teorema di Fréchet-Riesz

Situazione 8.1. U, V, W siano spazi di Hilbert su \mathbb{K} .

Teorema 8.2 (Fréchet-Riesz). Sia $\alpha \in V'$. Allora esiste un unico $y \in V$ tale che $\alpha(x) = \|x, y\|$ per ogni $x \in V$. Inoltre $\|\alpha\| = \|y\|$.

Dimostrazione. (1) L'unicità è immediata, perché se $\|x, y_1\| = \|x, y_2\|$ per ogni $x \in V$, allora $\|x, y_1 - y_2\| = 0$, per ogni $x \in V$ e quindi $y_1 - y_2 \in V^\perp \stackrel{7.4}{=} 0$.

(2) Sia $\alpha = \bigcirc_x \|x, y\|$. Dimostriamo che $\|\alpha\| = \|y\|$. Infatti

$$|\alpha(x)| = \|\|x, y\|\| \leq \|x\| \|y\| \text{ implica } \frac{|\alpha(x)|}{\|x\|} = \|\alpha\| \leq \|y\|, \text{ per cui da}$$

$$|\alpha(y)| = \|\|y, y\|\| = \|y\| \|y\| \text{ segue } \|\alpha\| = \|y\|.$$

(3) Sia $\ker \alpha = V$. Allora $\alpha = 0$ e possiamo prendere $y = 0$.

(4) Assumiamo quindi $\ker \alpha \neq V$. Dalla continuità di α segue che $\ker \alpha$ è un sottospazio chiuso di V , cosicché il cor. 7.23 implica che $(\ker \alpha)^\perp \neq 0$. Possiamo quindi scegliere un $z \in (\ker \alpha)^\perp$ con $z \neq 0$.

Allora $\alpha(z) \neq 0$; altrimenti si annulla $\|z, z\| = 0$, in contrasto con l'ipotesi $z \neq 0$

Sostituendo z con $z/\alpha(z)$ possiamo assumere che $\alpha(z) = 1$.

Sia $x \in V$.

(A) Dimostriamo che $x - \alpha(x)z \in \ker \alpha$. Infatti

$$\alpha(x - \alpha(x)z) = \alpha(x) - \alpha(x)\alpha(z) = \alpha(x) - \alpha(x) = 0$$

(B) Perciò $0 = \|x - \alpha(x)z, z\| = \|x, z\| - \alpha(x)\|z\|^2$, per cui $\|x, z\| = \alpha(x)\|z\|^2$.

Poniamo $y := \frac{z}{\|z\|^2}$. Allora

$$\|x, y\| = \frac{1}{\|z\|^2} \|x, z\| = \frac{1}{\|z\|^2} \alpha(x) \|z\|^2 = \alpha(x)$$

Osservazione 8.3. La portata di questo risultato apparentemente astratto si vede dal caso $V = L^2$ (cfr. def. 1.27).

Sia $\alpha : L^2 \rightarrow \mathbb{K}$ un'applicazione lineare e continua. Per il teorema 8.2 allora esiste un $g \in L^2$ tale che per ogni $f \in L^2$ si ha

$$\alpha(f) := \int f(\omega) \overline{g(\omega)} d\omega$$

Osservazione 8.4. Sia $y \in V$. Allora $\alpha := \bigcirc_x \|x, y\| : V \rightarrow \mathbb{K}$ è elemento di V' , come segue dalla sesquilinearità e dalla continuità del prodotto scalare. Per il teorema 8.2 l'applicazione

$$\begin{aligned} V &\xrightarrow{\theta} V' \\ y &\mapsto \bigcirc_x \|x, y\| \end{aligned}$$

è biiettiva. Denotando con $\# : V' \longrightarrow V$ la sua inversa, abbiamo

$$\alpha(x) = \|x, \#\alpha\|$$

per ogni $\alpha \in V'$ ed ogni $x \in V$.

Si noti che per $\lambda \in \mathbb{K}$ si ha $\#\lambda\alpha = \bar{\lambda}\#\alpha$.

Infatti $\|x, \#\lambda\alpha\| = \lambda\alpha(x) = \lambda\|x, \#\alpha\| = \|x, \bar{\lambda}\#\alpha\|$.

Proposizione 8.5. Sia $\mathbb{K} = \mathbb{R}$. Allora la biiezione naturale

$$\bigcirc_y \bigcirc_x \|x, y\| : V \xrightarrow{\theta} V'$$

è lineare e un'isometria nella norma, si ha cioè $\|\#\alpha\| = \|\alpha\|$ per ogni $\alpha \in V'$.

Dimostrazione. Ciò segue direttamente dal teorema 8.2 e dall'oss. 8.4. Si noti che l'applicazione $\bigcirc_y \bigcirc_x \|x, y\|$ non è più lineare nel caso

$\mathbb{K} = \mathbb{C}$ perché $\|x, \lambda y\| = \bar{\lambda}\|x, y\|$ per $\lambda \in \mathbb{C}$.

Definizione 8.6. Un'applicazione $\varphi : V \longrightarrow W$ tra spazi vettoriali su \mathbb{C} si dice *antilineare*, se per ogni $x, y \in V$ ed ogni $\lambda \in \mathbb{C}$ si ha

$$\varphi(x + y) = \varphi x + \varphi y$$

$$\varphi\lambda x = \bar{\lambda}\varphi x$$

Se φ è biiettiva, allora φ si chiama *antiisomorfismo*.

Proposizione 8.7. Sia $\mathbb{K} = \mathbb{C}$. Allora la biiezione naturale

$\bigcirc_y \bigcirc_x \|x, y\| : V \xrightarrow{\theta} V'$ è un'antiisomorfismo nella norma.

Dimostrazione. Ciò segue dal teorema 8.2 e dall'oss. 8.4.

Corollario 8.8. V' è uno spazio di Hilbert con il prodotto scalare

$$\|\alpha, \beta\| := \|\#\beta, \#\alpha\|.$$

La norma che corrisponde a questo prodotto scalare coincide con quella definita nella def. 4.5.

Osservazione 8.9. Nel caso $V = \mathbb{K}^n$ con il prodotto scalare euclideo

per $\alpha \in V'$ il vettore $y := \#\alpha \in \mathbb{K}^n$ è dato da $y = \|x, y\| = \begin{pmatrix} x^1 \bar{y}^1 \\ \vdots \\ x^n \bar{y}^n \end{pmatrix}$.

Dimostrazione. Per $x = x^1 \delta_1 + \dots + x^n \delta_n \in \mathbb{K}^n$ abbiamo $\alpha(x) = x^1 \alpha(\delta_1) + \dots + x^n \alpha(\delta_n) = x^1 \bar{y}^1 + \dots + x^n \bar{y}^n = \|x, y\|$.

Lemma 8.10. Sia $\varphi \in \mathcal{L}(V, W)$. Allora per $w \in W$ l'applicazione $\varphi_w := \bigcirc_v \|\varphi v, w\| : V \longrightarrow \mathbb{K}$ è evidentemente lineare e continua e quindi si ha $\varphi_w \in V'$. Inoltre $\|\varphi_w\| \leq \|\varphi\| \|w\|$.

Dimostrazione. Dobbiamo dimostrare solo che $\|\varphi_W\| \leq \|\varphi\|\|w\|$.

Ma per $v \in V$ abbiamo $|\varphi_W(v)| = \|\varphi v, w\| \leq \|\varphi v\|\|w\| \leq \|\varphi\|\|v\|\|w\|$ e ciò implica $\|\varphi w\| \leq \|\varphi\|\|w\|$.

Teorema 8.11. Sia $\varphi \in \mathcal{L}(V, W)$. Allora esiste, univocamente determinato, un elemento $\varphi^* \in \mathcal{L}(W, V)$ tale che $\|\varphi v, w\| = \|v, \varphi^* w\|$ per ogni $v \in V$ e ed ogni $w \in W$.

Esplicitamente per ogni $w \in W$ si ha $\varphi^* w = \# \bigcirc_v \|\varphi v, w\|$.

Inoltre $\|\varphi^*\| \leq \|\varphi\|$ (nella prop. 8.13 vedremo che $\|\varphi^*\| = \|\varphi\|$).

Dimostrazione. Sia $w \in W$.

(1) Come nel lemma 8.10 consideriamo l'applicazione

$\varphi_w := \bigcirc_v \|\varphi v, w\| \in V'$. Per il teorema 8.2 è perciò definita l'applicazione $\varphi^* := \bigcirc_w \# \varphi_w : W \rightarrow V$. Quest' applicazione è lineare, infatti l'additività è evidente, mentre per $\lambda \in \mathbb{K}$ si ha

$$\varphi_{\lambda w} = \bigcirc_v \|\varphi v, \lambda w\| = \bigcirc_v \bar{\lambda} \|\varphi v, w\| = \bar{\lambda} \varphi_w \text{ e quindi}$$

$$\varphi^* \lambda w = \# \varphi_{\lambda w} = \# \lambda \varphi_w = \lambda \# \varphi_w = \lambda \varphi^* w,$$

usando l'osservazione 8.4.

(2) Dalla definizione abbiamo $\|v, \varphi^* w\| = \|\varphi v, w\|$ per ogni $v \in V$.

(3) Dal teorema 8.2 segue anche l'unicità di φ^* .

(4) Infine $\|\varphi^* w\| \stackrel{8.2}{=} \|\varphi w\| \stackrel{8.10}{\leq} \|\varphi\|\|w\|$. Ciò mostra che $\|\varphi^*\| \leq \|\varphi\| < \infty$ e quindi anche la continuità di φ^* .

Definizione 8.12. Nella situazione dell'oss. 8.11 l'operatore φ^* si chiama l'operatore aggiunto di φ . Otteniamo in questo modo un'applicazione

$$* : \mathcal{L}(V, W) \rightarrow \mathcal{L}(W, V)$$

e quindi, nel caso particolare $V = W$, un'applicazione

$$* : \mathcal{L}(V) \rightarrow \mathcal{L}(V).$$

Proposizione 8.13. Usando il simbolo $*$ sia nella direzione $\mathcal{L}(V, W) \rightarrow \mathcal{L}(W, V)$ che in quella opposta, per $\varphi, \psi \in \mathcal{L}(V, W)$ e $\lambda \in \mathbb{K}$ abbiamo le seguenti relazioni:

$$(1) (\varphi + \psi)^* = \varphi^* + \psi^*.$$

$$(2) (\lambda \varphi)^* = \bar{\lambda} \varphi^*.$$

$$(3) \varphi^{**} = \varphi.$$

$$(4) \|\varphi^*\| = \|\varphi\|.$$

Dimostrazione. Usando l'unicità dell'operatore aggiunto è sufficiente verificare le proprietà che lo caratterizzano. Siano $v \in V$ e $w \in W$.

Allora:

$$\begin{aligned}
 (1) \quad \|\varphi + \psi\| v, w \| &= \|\varphi v + \psi v, w\| = \|\varphi v, w\| + \|\psi v, w\| \\
 &= \|v, \varphi^* w\| + \|v, \psi^* w\| = \|v, \varphi^* v + \psi^* w\| \\
 &= \|v, (\varphi^* + \psi^*) w\|
 \end{aligned}$$

$$(2) \quad \|\lambda \varphi v, w\| = \lambda \|\varphi v, w\| = \lambda \|v, \varphi^* w\| = \|v, \bar{\lambda} \varphi^* w\|.$$

$$(3) \quad \|v, \varphi^* w\| = \|\varphi v, w\| \text{ significa proprio che } \varphi^{**} = \varphi.$$

(4) Per il teorema 8.11 abbiamo anche

$$\|\varphi\| = \|\varphi^{**}\| \leq \|\varphi^*\| \leq \|\varphi\|$$

Proposizione 8.14.

(1) Siano $\varphi \in \mathcal{L}(U, V)$ e $\psi \in \mathcal{L}(V, W)$. Allora $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.

(2) Se $\varphi \in \mathcal{L}(V, W)$ è biiettiva, allora $\varphi^{-1} \in \mathcal{L}(W, V)$ e si ha $(\varphi^{-1})^* = (\varphi^*)^{-1}$.

Dimostrazione. (1) Per $u \in U$ e $w \in W$ abbiamo

$$\|\psi \varphi u, w\| = \|\varphi u, \psi^* w\| = \|u, \varphi^* \psi^* w\|$$

(2) Per un teorema profondo e fondamentale dell'analisi funzionale anche l'inversa φ^{-1} è continua (cfr. Heuser, pag. 243). Dal punto (1) abbiamo adesso $\text{id} = \text{id}^* = (\varphi \circ \varphi^{-1})^* = (\varphi^{-1})^* \circ \varphi^*$.

Osservazione 8.15. Se con $j := \bigcirc_x \bigcirc_y \alpha(x) : V \rightarrow V''$ denotiamo

l'immersione canonica di uno spazio normato V nel proprio *biduale* V'' e con $\theta' := \bigcirc_\alpha \bigcirc_\beta \|\beta, \alpha\| : V' \rightarrow V''$ la biiezione canonica, allora

$$j = \theta' \circ \theta.$$

Ciò mostra che (nel caso di uno spazio di Hilbert) j è un isomorfismo, cosicché vediamo che ogni spazio di Hilbert è *riflessivo*. Per i dettagli cfr. Heuser, pagg. 324-326, 336-339.

Dimostrazione. Dobbiamo solo dimostrare che $j = \theta' \circ \theta$. Siano $x \in V$ ed $\alpha \in V'$. Allora

$$(\theta' \theta x)(\alpha) = \|\alpha, \theta x\| = \|\# \theta x, \# \alpha\| = \|x, \# \alpha\| = \alpha(x)$$

9. Il rapporto di Rayleigh

Situazione 9.1. Siano V uno spazio di Hilbert, $\varphi \in \mathcal{L}(V)$ e $v \in V \setminus 0$.

Definizione 9.2. Il numero $\varphi : v := \frac{\|\varphi v, v\|}{\|v, v\|}$ si chiama il *rapporto di Rayleigh* di φ in v .

Si noti che $\varphi : v$ dipende anche da $\| \cdot \|$.

Nota 9.3. e_1, \dots, e_s siano elementi di $V \setminus 0$ tra loro ortogonali. Sia $W := \text{SV}(e_1, \dots, e_s)$. Per l'oss. 7.19 W è chiuso in V . Per $x \in V$ la proiezione ortogonale di x su W è data da

$$p = \alpha_1 e_1 + \dots + \alpha_s e_s$$

con $\alpha_k = \frac{\|v, e_k\|}{\|e_k, e_k\|}$ per ogni k .

Dimostrazione. $v - p$ deve essere ortogonale ad e_k per ogni k e quindi deve valere $0 = \|v - p, e_k\|$, ovvero $\|v, e_k\| = \|p, e_k\|$. Per l'ortogonalità degli e_j abbiamo però

$$\|p, e_k\| = \|\alpha_k e_k, e_k\| = \alpha_k \|e_k, e_k\|, \text{ cosicch\`e } \alpha_k \|e_k, e_k\| = \|v, e_k\|.$$

Osservazione 9.4. Vediamo in particolare che il rapporto di Rayleigh $\varphi : v$ non è altro che la coordinata rispetto alla base $\{v\}$ della proiezione ortogonale di φv sulla retta $\mathbb{R}v$. Ciò implica che $\varphi v - (\varphi : v)v \perp v$.

Corollario 9.5. $\|\varphi v\| \geq |\varphi : v| \|v\|$.

Dimostrazione. Per il teorema di Pitagora (prop. 4.18) e usando l'oss. 9.4 abbiamo

$$\|\varphi v\|^2 = \|\varphi v - (\varphi : v)v\|^2 + \|(\varphi : v)v\|^2$$

e quindi $\|\varphi v\| \geq \|(\varphi : v)v\| = |\varphi : v| \|v\|$.

In alternativa l'enunciato può essere dedotto dalla disuguglianza di Cauchy-Schwarz:

$$|\varphi : v| = \frac{\|\varphi v, v\|}{\|v, v\|} \leq \frac{\|\varphi v\| \|v\|}{\|v\|^2} \leq \frac{\|\varphi v\|}{\|v\|}$$

Definizione 9.6. $\frac{\|\varphi v - (\varphi : v)v\|}{\|v\|}$ si chiama il *raggio di Weinstein* di φ in v .

Corollario 9.7. $\varphi : v = 0 \iff \varphi v \perp v$.

Dimostrazione. Segue dalla def. 9.2.

Teorema 9.8. φ sia simmetrico. Allora

$$\|\varphi\| = \sup \{|\varphi : x| \mid x \in (\|V\| = 1)\}$$

Dimostrazione. (1) Si osservi che per $x \in (\|V\| = 1)$ si ha $\varphi : x = \|\varphi x, x\|$, per cui $|\varphi : x| \leq \|\varphi x\| \|x\| = \|\varphi x\| \leq \|\varphi\|$, cosicch\`e vediamo che

$$A := \sup \{|\varphi : x| \mid x \in (\|V\| = 1)\} \leq \|\varphi\| < \infty.$$

(2) Seguiamo Schröder, pag. 39. Osservando che

$$\begin{aligned} \text{(i)} \quad & \|\varphi(x+y), x+y\| - \|\varphi(x-y), x-y\| = \\ & = 2\|\varphi x, y\| + 2\|\varphi y, x\| = 4 \operatorname{Re} \|\varphi x, y\| \end{aligned}$$

$$\text{(ii)} \quad \|\varphi x, x\| = \left\| \left\| \varphi \frac{x}{\|x\|}, \frac{x}{\|x\|} \right\| \right\| \|x\|^2 \leq |\varphi : x| \|x\|^2$$

per $x \neq 0$, per $\|x\| = \|y\| = 1$ abbiamo

$$4 \operatorname{Re} \|\varphi x, y\| \leq A (\|x+y\|^2 - \|x-y\|^2) = 2A (\|x\|^2 + \|y\|^2) = 4A$$

Se scriviamo $\|\varphi x, y\| = e^{i\lambda} \|\varphi x, y\|$ e sostituiamo x con $e^{i\lambda}x$, nell'ultima disuguaglianza otteniamo $\|\varphi x, y\| \leq A$ quindi anche $\|\varphi^*y\| \leq A$.

Dalla prop. 8.13 concludiamo che $\|\varphi\| = \|\varphi^*\| \leq A$.

10. Sistemi ortonormali completi

Situazione 10.1. Sia V uno spazio di Hilbert. Quando non indicato diversamente, I è un insieme non vuoto qualsiasi (di indici).

Definizione 10.2. Una famiglia $\bigcirc_{i \in I} X_i$ di elementi di V si dice

- (1) *ortogonale*, se $x_i \perp x_j$ per ogni $i, j \in I$ con $i \neq j$;
- (2) *ortonormale*, se è ortogonale e $\|x_i\| = 1$ per ogni $i \in I$.

Il concetto di insieme ortogonale è stato introdotto nella def. 4.20.

Definizione 10.3. Una famiglia $\bigcirc_{i \in I}$ di elementi di V si dice

(incondizionatamente) *sommabile*, se esiste $x \in V$ tale che per ogni $\varepsilon > 0$ esiste un sottoinsieme finito F di I tale che per ogni sottoinsieme finito E di I si abbia $\left\| x - \sum_{i \in E \cup F} x_i \right\| < \varepsilon$.

Si dimostra facilmente che allora x è univocamente determinato, per cui possiamo scrivere $x = \sum_{i \in I} x_i$.

Lemma 10.4. Una famiglia $\bigcirc_{i \in I} x_i$ di elementi di V è sommabile se e solo se per ogni $\varepsilon > 0$ esiste un sottoinsieme finito F di I tale che per ogni sottoinsieme finito H di I con $H \cap F = \emptyset$ si abbia $\left\| \sum_{h \in H} x_h \right\| < \varepsilon$.

Dimostrazione. (1) $\bigcirc_{i \in I} x_i$ sia sommabile con $\sum_{i \in I} x_i = x$ ed $\varepsilon > 0$. Per ipotesi esiste un sottoinsieme finito F di I tale che per ogni sottoinsieme finito E di I si abbia $\left\| x - \sum_{i \in E \cup F} x_i \right\| < \varepsilon$. Sia H un sottoinsieme finito di I con $H \cap F = \emptyset$. Allora

$$\begin{aligned} \left\| \sum_{h \in H} x_h \right\| &= \left\| \sum_{i \in H \cup F} x_i - \sum_{j \in F} x_j \right\| \leq \left\| \sum_{i \in H \cup F} x_i - x \right\| + \left\| x - \sum_{j \in F} x_j \right\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

(2) La condizione dell'enunciato sia soddisfatta.

Allora per ogni $n \in \mathbb{N} + 1$ esiste un sottoinsieme finito F_n di I tale che per ogni sottoinsieme finito H di I con $H \cap F_n = \emptyset$ vale $\left\| \sum_{h \in H} x_h \right\| < \frac{1}{n}$.

Ciò vale anche se sostituiamo F_n con $F_1 \cup \dots \cup F_n$, per cui possiamo assumere che $F_1 \subset F_2 \subset F_3 \subset \dots$

Per $n \in \mathbb{N} + 1$ poniamo $y_n := \sum_{i \in F_n} x_i$. Siano $\varepsilon > 0$ ed n tali che $\frac{1}{n} < \varepsilon$.

Per $k \geq l \geq n$ allora

$$\|y_k - y_l\| = \left\| \sum_{i \in F_k \setminus F_l} x_i \right\| < \frac{1}{l} \leq \frac{1}{n} < \varepsilon$$

e vediamo che $\bigcirc_n y_n$ è una successione di Cauchy. Siccome lo spazio di Hilbert V è completo, esiste $x \in V$ tale che $\bigcirc_n y_n \rightarrow x$. Dimostriamo che $x = \sum_{i \in I} x_i$. Sia $\varepsilon > 0$. Allora esiste $n \in \mathbb{N} + 1$ tale che allo stesso tempo $\frac{1}{n} < \frac{\varepsilon}{2}$ e $\|x - y_n\| < \frac{\varepsilon}{2}$ per ogni $m \geq n$. Siano F_n come prima ed E un sottoinsieme finito di I . Allora

$$\begin{aligned} \left\| x - \sum_{i \in F_n \cup E} x_i \right\| &= \left\| x - \sum_{i \in F_n} x_i - \sum_{h \in E \setminus F_n} x_h \right\| \\ &= \|x - y_n\| + \left\| \sum_{h \in E \setminus F_n} x_h \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Osservazione 10.5. La famiglia $\bigcirc_{i \in I} x_i$ sia sommabile e $\sigma : I \rightarrow I$ un'applicazione biiettiva. Allora anche la famiglia $\bigcirc_{i \in I} x_{\sigma(i)}$ è sommabile.

Dimostrazione. Siano $x := \sum_{i \in I} x_i$ ed $\varepsilon > 0$. Scegliamo F come nella def. 10.3. Sia E un sottoinsieme finito di I . Allora

$\{\sigma(i) \mid i \in \sigma^{-1}(F) \cup E\} = \sigma(\sigma^{-1}(F) \cup E) = F \cup \sigma(E)$, per cui

$$\left\| x - \sum_{i \in \sigma^{-1}(F) \cup E} x_{\sigma(i)} \right\| = \left\| x - \sum_{i \in F \cup \sigma(E)} x_i \right\| < \varepsilon$$

Lemma 10.6. Sia $\bigcirc_{i \in I} x_i$ una famiglia di elementi di V ed $x \in V$. Allora sono equivalenti:

(1) La famiglia è sommabile con $\sum_{i \in I} x_i = x$.

(2) L'insieme $J := \{i \in I \mid x_i \neq 0\}$ è finito oppure $|J| = |\mathbb{N}|$ con $J = \{i_1, i_2, \dots\}$ e in questo secondo caso per ogni permutazione

$\sigma : \mathbb{N} \rightarrow \mathbb{N}$ si ha $x = \lim_{n \rightarrow \infty} \sum_{k=1}^n x_{i_{\sigma(k)}}$.

Dimostrazione. (1) \implies (2): Come nella dimostrazione della prop. 10.4 per ogni $n \in \mathbb{N} + 1$ esiste un sottoinsieme finito F_n di I tale che per ogni sottoinsieme finito H di I con $H \cap F_n = \emptyset$ vale $\left\| \sum_{h \in H} x_h \right\| < \frac{1}{n}$.

L'insieme $K := \bigcup_{n=1}^{\infty} F_n$ è numerabile e inoltre $J \subset K$. Infatti per

$i \in I \setminus K$ e per ogni $n \in \mathbb{N} + 1$ abbiamo $\|x_i\| < \frac{1}{n}$ (prendendo $H := \{i\}$ nelle considerazioni precedenti) e quindi $x_i = 0$. Il resto è adesso evidente.

(2) \implies (1): Chiaro.

Lemma 10.7. Siano $\bigcirc_{i \in I} x_i$ e $\bigcirc_{i \in I} y_i$ famiglie sommabili di elementi di X con lo stesso insieme di indici I . Siano $\lambda \in \mathbb{K}$ e $z \in X$. Allora:

$$(1) \sum_{i \in I} x_i + \sum_{i \in I} y_i = \sum_{i \in I} (x_i + y_i).$$

$$(2) \lambda \sum_{i \in I} x_i = \sum_{i \in I} \lambda x_i.$$

$$(3) \left\| \sum_{i \in I} x_i, z \right\| = \sum_{i \in I} \|x_i, z\|.$$

Dimostrazione. (1) Chiaro.

(2) Chiaro.

(3) Per il lemma 10.6 possiamo assumere che $I = \mathbb{N}$. Siano $x := \sum_{i=0}^{\infty} x_i$ ed $\varepsilon > 0$. Allora esiste $m \in \mathbb{N}$ tale che $\left\| \sum_{i=m+1}^{\infty} x_i \right\| \leq \varepsilon$, per cui dalla disuguaglianza di Cauchy-Schwarz otteniamo

$$\begin{aligned} \left| \|x, z\| - \sum_{i=0}^m \|x_i, z\| \right| &= \left| \sum_{i=0}^{\infty} \|x_i, z\| - \sum_{i=0}^m \|x_i, z\| \right| \\ &= \left| \sum_{i=m+1}^{\infty} \|x_i, z\| \right| \leq \left\| \sum_{i=m+1}^{\infty} x_i \right\| \|z\| \leq \varepsilon \|z\| \end{aligned}$$

Lemma 10.8. Sia $\bigcirc_{i \in I} x_i$ una famiglia ortogonale di elementi di X . Allora sono equivalenti:

(1) La famiglia $\bigcirc_{i \in I} x_i$ è sommabile in X .

(2) La famiglia $\bigcirc_{i \in I} \|x_i\|^2$ è sommabile in \mathbb{R} .

In tal caso inoltre $\left\| \sum_{i \in I} x_i \right\|^2 = \sum_{i \in I} \|x_i\|^2$.

Dimostrazione. Per il teorema di Pitagora per ogni sottoinsieme finito G di I abbiamo $\sum_{g \in G} \|x_g\|^2 = \left\| \sum_{g \in G} x_g \right\|^2$. È chiaro adesso che (1) e (2) sono equivalenti.

La famiglia $\bigcirc_{i \in I} x_i$ sia sommabile. Allora dal lemma 10.7 e sfruttando l'ipotesi di ortogonalità otteniamo

$$\left\| \sum_{i \in I} x_i \right\|^2 = \left\| \sum_{i \in I} x_i, \sum_{i \in I} x_i \right\| = \sum_{i \in I} \|x_i, x_i\| = \sum_{i \in I} \|x_i\|^2$$

Lemma 10.9. Siano e_1, \dots, e_m elementi ortonormali di V ed $x \in V$. Allora:

$$(1) \left\| x - \sum_{j=1}^m \|x, e_j\| e_j \right\|^2 = \|x\|^2 - \sum_{j=1}^m \|x, e_j\|^2.$$

$$(2) \sum_{j=1}^m \|x, e_j\|^2 \leq \|x\|^2.$$

Dimostrazione. (1) Usando l'ortonormalità di e_1, \dots, e_m abbiamo

$$\begin{aligned} \left\| x - \sum_{j=1}^m \|x, e_j\| e_j \right\|^2 &= \|x\|^2 - 2 \operatorname{Re} \sum_{j=1}^m \|x, e_j\|^2 + \sum_{j=1}^m \|x, e_j\|^2 \\ &= \|x\|^2 - \sum_{j=1}^m \|x, e_j\|^2 \end{aligned}$$

(2) Ciò segue dal punto (1) perché sicuramente

$$\left\| x - \sum_{j=1}^m \|x, e_j\| e_j \right\|^2 \geq 0.$$

Lemma 10.10. Siano $\bigcirc_{i \in I} \alpha_i$ una famiglia di numeri reali ≥ 0 ed $a \in \mathbb{R}$ tali che per ogni sottoinsieme finito F di I si abbia $\sum_{j \in F} \alpha_j \leq a$.

Allora la famiglia $\bigcirc_{i \in I} \alpha_i$ è sommabile e si ha $\sum_{i \in I} \alpha_i \leq a$.

Dimostrazione. Per il lemma 10.6 e noti teoremi sulle serie numeriche con coefficienti non negativi è sufficiente dimostrare che l'insieme

$G := \{j \in I \mid \alpha_j \neq 0\}$ è numerabile. Però $G := \bigcup_{n=1}^{\infty} G_n$ dove

$G_n := \left\{ j \in I \mid \alpha_j \geq \frac{1}{n} \right\}$ e l'ipotesi nell'enunciato implica che ogni G_n è finito. Perciò G è sicuramente numerabile.

Proposizione 10.11. Siano $\bigcirc_{i \in I} e_i$ una famiglia ortonormale di elementi di V ed $x \in V$. Allora:

$$(1) \sum_{i \in I} \|x, e_i\|^2 \leq \|x\|^2 \text{ (disuguaglianza di Bessel).}$$

$$(2) \sum_{i \in I} \|x, e_i\|^2 = \|x\|^2 \text{ se e solo se } x = \sum_{i \in I} \|x, e_i\| e_i \text{ (identità di Parseval).}$$

Dimostrazione. (1) Sia F un sottoinsieme finito di I . Per il lemma 10.9 abbiamo $\sum_{i \in F} \|x, e_i\|^2 \leq \|x\|^2$. Dal lemma 10.9 segue che la somma

$\sum_{i \in I} \|x, e_i\|^2$ esiste e che è soddisfatta la disuguaglianza di Bessel.

(2) Le seguenti affermazioni sono equivalenti:

$$x = \sum_{i \in I} \|x, e_i\| e_i$$

$$\left\| x - \sum_{i \in I} \|x, e_i\| e_i \right\|^2 = 0$$

$$\|x\|^2 = \sum_{i \in I} \|x, e_i\|^2$$

Qui, nell'equivalenza tra l'ultima e la penultima uguaglianza, abbiamo usato ancora il lemma 10.9.

Corollario 10.12. Siano $\bigcirc_{i \in I} e_i$ una famiglia ortonormale in V ed $x \in V$.

Allora $y := \sum_{i \in I} \|x, e_i\| e_i$ esiste e si ha $\|y\|^2 = \sum_{i \in I} \|x, e_i\|^2$.

Dimostrazione. Per $i \in I$ sia $x_i := \|x, e_i\| e_i$. Allora $\bigcirc_{i \in I} x_i$ è una famiglia ortogonale in V a cui possiamo applicare il lemma 10.8. Notiamo che $\|x_i\|^2 = \|x, e_i\|^2$, cosicché dalla prop. 10.11 segue che la famiglia $\bigcirc_{i \in I} \|x_i\|^2$ è sommabile.

Per il lemma 10.8 però $\sum_{i \in I} x_i = \sum_{i \in I} \|x, e_i\| e_i$ esiste e in più

$$\|y\|^2 = \sum_{i \in I} \|x, e_i\|^2.$$

Definizione 10.13. Una base di Hilbert di V è una famiglia ortonormale $\bigcirc_{i \in I} e_i$ di elementi di V tale che per ogni $x \in V$ valga lo sviluppo di Fourier $x = \sum_{i \in I} \|x, e_i\| e_i$.

Teorema 10.14. Sia $\bigcirc_{i \in I} e_i$ una famiglia ortonormale di elementi di V . Allora sono equivalenti:

- (1) $\bigcirc_{i \in I} e_i$ è una base di Hilbert di V .
- (2) $\{e_i \mid i \in I\}^\perp = 0$.
- (3) $\{e_i \mid i \in I\}$ è un sistema ortonormale massimale.

Dimostrazione. (1) \implies (2): $\bigcirc_{i \in I} e_i$ sia una base di Hilbert ed $x \perp e_i$ per ogni $i \in I$. Per ipotesi allora $x = \sum_{i \in I} \|x, e_i\| e_i = \sum_{i \in I} 0 \cdot e_i = 0$.

(2) \implies (3): Chiaro perché nell'ipotesi (2) non possiamo aggiungere un elemento alla famiglia data che sia $\neq 0$, ma ortogonale a tutti gli elementi della famiglia.

(3) \implies (1): Sia $x \in V$. Per il cor. 10.12 esiste $y := \sum_{i \in I} \|x, e_i\| e_i$. Dobbiamo dimostrare che $y = x$. Per ogni $j \in I$ abbiamo

$$\|y - x, e_j\| = \|y, e_j\| - \|x, e_j\| = \| \|x, e_j\| e_j, e_j \| - \|x, e_j\| = 0$$

cosicché per l'ipotesi di massimalità necessariamente $y - x = 0$.

Proposizione 10.15. Sia $\bigcirc_{i \in I} e_i$ una famiglia ortonormale di elementi di V . Allora sono equivalenti:

- (1) $\bigcirc_{i \in I} e_i$ è una base di Hilbert di V .

(2) Per ogni $x, y \in V$ vale $\|x, y\| = \sum_{i \in I} \|x, e_i\| \|e_i, y\|$.

(3) Per ogni $x \in V$ vale $\|x\|^2 = \sum_{i \in I} \|x, e_i\|^2$.

(4) $\overline{\text{SV}(\{e_i \mid i \in I\})} = V$.

Dimostrazione. Verifica immediata, utilizzando il teorema 10.14.

Proposizione 10.16. Siano $\bigcirc_{i \in I} e_i$ e $\bigcirc_{j \in J} d_j$ due basi di Hilbert di V .

Allora $|I| = |J|$.

Dimostrazione. Ciò è chiaro se uno dei due insiemi I e J è finito. Nel caso generale per ogni $i \in I$ l'insieme $K_i := \{j \in J \mid \|e_i, d_j\| \neq 0\}$ è numerabile per i risultati precedenti. D'altra parte $J := \bigcup_{i \in I} K_i$.

Infatti per ogni $j \in J$ deve esistere un $i \in I$ con $\|e_i, d_j\| \neq 0$ perché altrimenti si avrebbe $d_j = 0$, come segue dal teorema 10.14. Con lo stesso ragionamento si vede che $K_i \neq \emptyset$ per ogni i . Se I e J sono infiniti, ciò implica che $|I| = |J|$.

Proposizione 10.17. V possiede una base di Hilbert.

Dimostrazione. Ciò segue dal lemma di Zorn, tenendo conto del teorema 10.14.

Definizione 10.18. La *dimensione di Hilbert* $\text{Dim } H$ di V è la cardinalità di una base di Hilbert di V .

Per le prop. 10.16 e 10.17 $\text{Dim } V$ è ben definita. È chiaro che $\dim V = \text{Dim } V$, se $\dim V < \infty$.

Proposizione 10.19. W sia uno spazio di Hilbert della stessa dimensione di Hilbert di V . Allora V e W sono isometricamente isomorfi.

Dimostrazione. Possiamo assumere che con lo stesso insiemi di indici $\bigcirc_{i \in I} e_i$ sia una base di Hilbert di V e $\bigcirc_{i \in I} d_i$ una base di Hilbert di W .

Dai risultati noti finora segue facilmente che tramite

$\sum_{i \in I} \alpha_i e_i \mapsto \sum_{i \in I} \alpha_i d_i$ otteniamo un isomorfismo isometrico $V \rightarrow W$.

11. Ortogonalizzazione

Situazione 11.1. Sia V uno spazio di Hilbert.

Osservazione 11.2. Siano e_1, \dots, e_m elementi ortonormali di V ed $x \in V$. Allora la proiezione ortogonale di x su $SV(e_1, \dots, e_m)$ è data da

$$p = \sum_{i=1}^m \|x, e_i\| e_i.$$

Dimostrazione. Ciò è un caso particolare della nota 9.3.

Osservazione 11.3. È noto (una dimostrazione si trova in Heuser, pagg. 159-161) che nella situazione dell'oss. 11.2 p è il vettore (univocamente determinato) con distanza minima da x in $SV(e_1, \dots, e_m)$.

Nota 11.4 (ortonormalizzazione di Schmidt). v_1, v_2, \dots siano vettori linearmente indipendenti di V . Troviamo un sistema ortonormale numerabile e_1, e_2, \dots nel modo seguente:

(1) $e_1 := \frac{v_1}{|v_1|}$ e poniamo $E_1 := SV(e_1) = SV(v_1)$

(2) p_2 sia la proiezione ortogonale di v_2 su E_1 . Per l'ipotesi di indipendenza lineare dei v_j sicuramente $v_2 \notin E_1$ si ha $v_2 - p_2 \neq 0$.

Poniamo $e_2 := \frac{v_2 - p_2}{|v_2 - p_2|}$ e $E_2 := SV(e_1, e_2) = SV(v_1, v_2)$.

(3) p_3 sia la proiezione ortogonale di v_3 su E_2 . Come prima vediamo che $v_3 - p_3 \neq 0$ e poniamo $e_3 := \frac{v_3 - p_3}{|v_3 - p_3|}$ e $E_3 := SV(e_1, e_2, e_3) = SV(v_1, v_2, v_3)$.

(4) In questo modo troviamo i vettori e_1, e_2, e_3, \dots . Per ogni $s = 1, 2, \dots$ si ha $SV(e_1, \dots, e_s) = SV(v_1, \dots, v_s)$, e ciò implica che anche $SV(e_1, e_2, \dots) = SV(v_1, v_2, \dots)$. I vettori e_i possono essere determinati algebricamente, infatti per l'oss. 11.2 abbiamo le relazioni

$$p_2 := \|v_2, e_1\| e_1$$

$$p_3 := \|v_3, e_1\| e_1 + \|v_3, e_2\| e_2$$

$$p_i := \|v_i, e_1\| e_1 + \dots + \|v_i, e_{i-1}\| e_{i-1}$$

⋮

$$p_{i+1} := \|v_{i+1}, e_1\| e_1 + \dots + \|v_{i+1}, e_i\| e_i$$

⋮

da cui troviamo che $e_i := \frac{v_i - p_i}{|v_i - p_i|}$

Proposizione 11.5. V sia separabile e $\dim V = \infty$. Allora $\dim V = |\mathbb{N}|$.

Dimostrazione. Per ipotesi esiste un insieme numerabile (necessariamente infinito) $M = \{x_1, x_2, \dots\}$ denso in V . Siano:

v_1 il primo (rispetto alla numerazione usata) elemento $\neq 0$ di M ,

v_2 il primo elemento di M non contenuto in $SV(v_1)$,

v_3 il primo elemento di M non contenuto in $SV(v_1, v_2)$, ecc...

Questi elementi esistono, perchè gli insiemi $V \setminus SV(v_1, \dots, v_s)$ sono tutti aperti e non vuoti di V ed M è, per ipotesi, denso in V .

Per costruzione gli elementi v_1, v_2, \dots sono linearmente indipendenti, inoltre $\overline{SV(v_1, v_2, \dots)} = V$.

Dalla nota 11.4 otteniamo un sistema ortonormale e_1, e_2, \dots tale che anche $\overline{SV(e_1, e_2, \dots)} = V$.

Dobbiamo dimostrare che $\bigcirc_n e_n$ è una base di Hilbert di V . Sia $x \in V$.

Siccome $\overline{SV(e_1, e_2, \dots)} = V$, esiste una successione $\bigcirc_n y_n \rightarrow x$ con

$$y_n = \sum_{k=1}^{m_n} \lambda_{n_k} e_k \text{ ed } m_1 < m_2 < \dots$$

Dal lemma 10.9 e dall'oss. 11.3 abbiamo

$$\begin{aligned} 0 &\leq \|x\|^2 - \sum_{k=1}^{m_n} \|x, e_k\|^2 = \left\| x - \sum_{k=1}^{m_n} \|x, e_k\| e_k \right\|^2 \\ &\leq \left\| x - \sum_{k=1}^{m_n} \lambda_{n_k} e_k \right\|^2 = \|x - y_n\|^2 \end{aligned}$$

cosicché necessariamente $\bigcirc_n \sum_{k=1}^{m_n} \|x, e_k\|^2 \rightarrow \|x\|^2$. Il cor. 10.12 e la

prop 10.15 implicano $x = \sum_{k=1}^{\infty} \|x, e_k\| e_k$.

12. Predizione lineare

Situazione 12.1. $\bigcirc_n X_n$ sia una STS $(0, \gamma)$ reale.

Identifichiamo, quando opportuno, X_n con il corrispondente elemento di L^2 .

Osservazione 12.2. Nelle ipotesi della situazione 12.1 per ogni $i, j \in \mathbb{Z}$ abbiamo

$$\begin{aligned} \text{cov}(X_i, X_j) &= \|X_i, X_j\| = \gamma(i - j) \\ &= \|X_i, X_j\| = \gamma(i - j) = \text{cov}(X_j, X_i) \end{aligned}$$

Per $k \in \mathbb{N}$ poniamo $\gamma_k := \gamma(k) = \gamma(-k)$.

Nota 12.3. Siano $n \in \mathbb{N}$ ed $Y \in L^2$.

Dalla prop. 7.36 sappiamo che $\text{SV}(X_0, \dots, X_{n-1})$ è chiuso in L^2 e quindi per il cor. 5.16 esiste, univocamente determinato, un punto $\hat{Y} \in W$ con $d(\hat{Y}, W) = \|Y - \hat{Y}\|$. Per il lemma 7.12 \hat{Y} è caratterizzato dalle condizioni $\hat{Y} \in W$ e $Y - \hat{Y} \perp W$.

Dobbiamo quindi trovare $a^0, a^1, \dots, a^{n-1} \in \mathbb{R}$ tali che

$$\|X_i, Y - a^0 X_0 - \dots - a^{n-1} X_{n-1}\| = 0 \quad (*)$$

per ogni $i = 0, \dots, n-1$, avendo allora $\hat{Y} = a^0 X_0 + \dots + a^{n-1} X_{n-1}$.

Le equazioni (*) si chiamano *equazioni normali* e costituiscono un sistema lineare

$$\begin{aligned} \|X_0, X_0\|a^0 + \|X_0, X_1\|a^1 + \dots + \|X_0, X_{n-1}\|a^{n-1} &= \|X_0, Y\| \\ \|X_1, X_0\|a^0 + \|X_1, X_1\|a^1 + \dots + \|X_1, X_{n-1}\|a^{n-1} &= \|X_1, Y\| \\ &\vdots \\ \|X_{n-1}, X_0\|a^0 + \|X_{n-1}, X_1\|a^1 + \dots + \|X_{n-1}, X_{n-1}\|a^{n-1} &= \|X_{n-1}, Y\| \end{aligned}$$

ovvero

$$\begin{aligned} \gamma_0 a^0 + \gamma_1 a^1 + \dots + \gamma_{n-1} a^{n-1} &= \|X_0, Y\| \\ &\vdots \\ \gamma_{n-1} a^0 + \gamma_{n-2} a^1 + \dots + \gamma_0 a^{n-1} &= \|X_{n-1}, Y\| \end{aligned}$$

Se poniamo

$$\begin{aligned} (\Omega|n) &:= \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-2} & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & & \gamma_{n-2} \\ \vdots & & & & \vdots \\ \gamma_{n-2} & & & \gamma_0 & \gamma_1 \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_1 & \gamma_0 \end{pmatrix} \\ (a|n) &:= \begin{pmatrix} a^0 \\ \vdots \\ a^{n-1} \end{pmatrix} \quad (b|n) := \begin{pmatrix} \|X_0, Y\| \\ \vdots \\ \|X_{n-1}, Y\| \end{pmatrix} \end{aligned}$$

il sistema diventa $(\Omega|n)(a|n) = (b|n)$.

Se i vettori X_0, \dots, X_{n-1} sono linearmente indipendenti, la matrice $(\Omega|n)$ è simmetrica, positivamente definita e invertibile e si ha

$$(a|n) = (\Omega|n)^{-1} (b|n)$$

Nota 12.4. Nelle ipotesi e con le notazioni della nota 12.3 sia $Y = X_{n+h}$ con $h \in \mathbb{N}$. Allora

$$(b|n) = \begin{pmatrix} \gamma_{n+h} \\ \gamma_{n+h-1} \\ \vdots \\ \gamma_{h+1} \end{pmatrix}$$

Le considerazioni della nota 12.3 possono essere usate per ottenere \hat{X}_{n+h} come predizione lineare di X_{n+h} .

Nel caso particolare $h = 0$ abbiamo $(b|n) = \begin{pmatrix} \gamma_n \\ \gamma_{n-1} \\ \vdots \\ \gamma_1 \end{pmatrix}$

e il sistema da risolvere diventa della forma

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{pmatrix} \begin{pmatrix} a^0 \\ a^1 \\ \vdots \\ a^{n-1} \end{pmatrix} = \begin{pmatrix} \gamma_n \\ \gamma_{n-1} \\ \vdots \\ \gamma_1 \end{pmatrix}$$

e prende il nome di *equazione di Yule-Walker*.

Esempio 12.5. Per comprendere la struttura della matrice $(\Omega|n)$ consideriamo il caso $n = 5$:

$$\begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \gamma_1 & \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_2 & \gamma_1 & \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_3 & \gamma_2 & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_4 & \gamma_3 & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix}$$

Osservazione 12.6. Nel prossimo capitolo presenteremo algoritmi per il calcolo veloce di $(\Omega|n)^{-1}$.

13. Matrici di Toeplitz

Situazione 13.1. Siano dati i numeri reali $\gamma_0 = 1, \gamma_1, \dots, \gamma_{n-1}$.

Per $1 \leq k \leq n$ le matrici $(\Omega|k) \in \mathbb{R}_k^k$ (dette *matrici di Toeplitz*) siano definite nel modo seguente:

$$\begin{aligned} (\Omega|1) &:= (\gamma_0) \\ (\Omega|2) &:= \begin{pmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{pmatrix} \\ (\Omega|3) &:= \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix} \\ (\Omega|4) &:= \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_1 & \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_2 & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_3 & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix} \\ &\vdots \\ (\Omega|n) &:= \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2} \\ \vdots & \vdots & & \vdots \\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0 \end{pmatrix} \end{aligned}$$

Per costruzione le matrici $(\Omega|k)$ sono simmetriche. Assumiamo inoltre che la matrice $(\Omega|n)$ sia positivamente definita.

Quando non indicato diversamente, supponiamo $1 \leq k \leq n$.

Osservazione 13.2. Denotiamo con $(J|k)$ la matrice di permutazione

$$(J|k) := \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{pmatrix}$$

La j -esima colonna di $(J|k)$ è quindi data da $(J|k)_j = \delta_{k-j+1}$,

la i -esima riga $(J|k)^i = \delta^{k-i+1}$.

Per $v \in \mathbb{R}^k$ abbiamo $((J|k)v)^i = v^{k-i+1}$,

per $f \in \mathbb{R}_k$ abbiamo $((J|k)f)_j = f_{k-j+1}$.

Osservazione 13.3. $(J|k)^{-1} = (J|k)$.

Dimostrazione. Per ogni i, j abbiamo

$$((J|k)(J|k))^i_j = (J|k)^i_j (J|k)_j = \delta^{k-i+1} \delta_{k-j+1} = \delta_j^i.$$

Definizione 13.4. Poniamo

$$(\gamma|k) := \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_k \end{pmatrix}$$

$$\varepsilon(k) := \begin{pmatrix} \gamma_k \\ \vdots \\ \gamma_1 \end{pmatrix} = (J|k) (\gamma|k)$$

Osservazione 13.5.

- (1) La i -esima diagonale laterale superiore o inferiore di $(\Omega|k)$ è costante e tutti i suoi coefficienti sono uguali a γ_i .
- (2) Per $k < n$ si ha

$$(\Omega|k+1) = \begin{pmatrix} (\Omega|k) & (\varepsilon|k) \\ (\varepsilon|k)^t & 1 \end{pmatrix}$$

Osservazione 13.6. Le matrici $(\Omega|k)$ sono tutte positivamente definite e quindi in particolare invertibili.

Dimostrazione. Ciò segue dall'oss. 13.5 e da un noto teorema di teoria delle matrici simmetriche.

Definizione 13.7. Una matrice $A \in \mathbb{R}_k^k$ si dice *persimmetrica*, se coincide con la matrice che si ottiene tramite la riflessione all'antidiagonale che unisce A_k^1 con A_1^k , se abbiamo cioè

$$A = \begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_{k-1}^1 & A_k^1 \\ A_1^2 & A_2^2 & \cdots & A_{k-1}^2 & A_k^2 \\ \vdots & \vdots & & \vdots & \vdots \\ A_1^{k-1} & A_2^{k-1} & \cdots & A_{k-1}^{k-1} & A_k^{k-1} \\ A_1^k & A_2^k & \cdots & A_{k-1}^k & A_k^k \end{pmatrix}$$

$$= \begin{pmatrix} A_k^k & A_{k-1}^{k-1} & \cdots & A_k^2 & A_k^1 \\ A_{k-1}^k & A_{k-1}^{k-1} & \cdots & A_{k-1}^2 & A_{k-1}^1 \\ \vdots & \vdots & & \vdots & \vdots \\ A_2^k & A_2^{k-1} & \cdots & A_2^2 & A_2^1 \\ A_1^k & A_1^{k-1} & \cdots & A_1^2 & A_1^1 \end{pmatrix}$$

Ciò accade se e solo se $A_j^i = A_{k-i+1}^{k-j+1}$ per ogni i, j .

Osservazione 13.8. (1) Una matrice $A \in \mathbb{R}_k^k$ è persimmetrica se e solo se $A = (J|k) A^t (J|k)$.

(2) Una matrice simmetrica $A \in \mathbb{R}_k^k$ è quindi persimmetrica se e solo se $A (J|k) = (J|k) A$.

Dimostrazione. (1) Sia $A \in \mathbb{R}_k^k$. Per ogni i, j allora

$$\begin{aligned} ((J|k) A^t (J|k))_j^i &= (J|k)_j^i A^t (J|k)_j = \delta^{k-i+1} A^t \delta_{k-j+1} \\ &= (A^t)_{k-j+1}^{k-i+1} = A_{k-i+1}^{k-j+1} \end{aligned}$$

(2) Ciò segue dal punto (1) usando l'oss. 13.3.

Osservazione 13.9. Le matrici $(\Omega|k)$ e $(\Omega|k)^{-1}$ sono simmetriche e persimmetriche.

Dimostrazione. Abbiamo già osservato che la matrice $(\Omega|k)$ è simmetrica. Perciò se $(\Omega|k)^{-1}$ è simmetrica, $(\Omega|k)$ è persimmetrica per il punto (1) dell'oss. 13.4, cosicché $(\Omega|k)(J|k) = (J|k)(\Omega|k)$ e quindi $(J|k)(\Omega|k)^{-1} = (\Omega|k)^{-1}(J|k)$ usando l'oss. 13.3.

Definizione 13.10. Poniamo

$$\begin{aligned}(y|k) &:= (\Omega|k)^{-1}(\gamma|k) \\ (z|k) &:= (\Omega|k)^{-1}(\varepsilon|k)\end{aligned}$$

Lemma 13.11. $(z|k) = (J|k)(y|k)$.

Dimostrazione. Infatti

$$\begin{aligned}(J|k)(y|k) &= (J|k)(\Omega|k)^{-1}(\gamma|k) \stackrel{13.8}{=} (\Omega|k)^{-1}(J|k)(\gamma|k) \\ &= (\Omega|k)^{-1}(\varepsilon|k) = (z|k)\end{aligned}$$

Lemma 13.12. $1 - (\varepsilon|k)^t(z|k) \neq 0$ per $1 \leq k < n$.

Dimostrazione. Siano $T := \begin{pmatrix} \delta & -(z|k) \\ 0 & 1 \end{pmatrix}$ ed $\alpha := 1 - (\varepsilon|k)^t(z|k)$.

Allora T è invertibile e quindi lo è anche

$$\begin{aligned}T^t(\Omega|k+1)T &= \begin{pmatrix} \delta & 0 \\ -(z|k)^t & 1 \end{pmatrix} \begin{pmatrix} (\Omega|k) & (\varepsilon|k) \\ (\varepsilon|k)^t & 1 \end{pmatrix} \begin{pmatrix} \delta & -(z|k) \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \delta & 0 \\ -(z|k)^t & 1 \end{pmatrix} \begin{pmatrix} (\Omega|k) & -(\Omega|k)(z|k) + (\varepsilon|k) \\ (\varepsilon|k)^t & -(\varepsilon|k)^t(z|k) + 1 \end{pmatrix} \\ &= \begin{pmatrix} \delta & 0 \\ -(z|k)^t & 1 \end{pmatrix} \begin{pmatrix} (\Omega|k) & 0 \\ (\varepsilon|k)^t & \alpha \end{pmatrix} \\ &= \begin{pmatrix} (\Omega|k) & 0 \\ -(z|k)^t(\Omega|k) + (\varepsilon|k)^t & \alpha \end{pmatrix} = \begin{pmatrix} (\Omega|k) & 0 \\ 0 & \alpha \end{pmatrix}\end{aligned}$$

per cui necessariamente $\alpha \neq 0$.

Nota 13.13. Possiamo determinare $(y|k)$ tramite un algoritmo ricorsivo nel modo seguente

(1) Ovviamente $(y|1) = (\gamma|1) = \gamma_1$.

(2) Assumiamo di conoscere $(y|k)$ e quindi per il lemma 13.11 anche $(z|k)$. $(y|k+1)$ è allora la soluzione dell'equazione

$$(\Omega|k+1)(y|k+1) = (\gamma|k+1)$$

che possiamo scrivere nella forma

$$\begin{pmatrix} (\Omega|k) & (\varepsilon|k) \\ (\varepsilon|k)^t & 1 \end{pmatrix} \begin{pmatrix} v \\ \lambda \end{pmatrix} = \begin{pmatrix} (\gamma|k) \\ \gamma_{k+1} \end{pmatrix}$$

con v e λ da determinare. Ciò equivale a

$$(\Omega|k)v + \lambda(\varepsilon|k) = (\gamma|k)$$

$$(\varepsilon|k)^t v + \lambda = \gamma_{k+1}$$

ovvero

$$v = (\Omega|k)^{-1}(\gamma|k) - \lambda(\Omega|k)^{-1}(\varepsilon|k) = (y|k) - \lambda(z|k)$$

$$\lambda = \gamma_{k+1} - (\varepsilon|k)^t v$$

da cui troviamo

$$\lambda = \gamma_{k+1} - (\varepsilon|k)^t (y|k) + \lambda(\varepsilon|k)^t (z|k)$$

cosicché $\lambda(1 - (\varepsilon|k)^t (z|k)) = \gamma_{k+1} - (\varepsilon|k)^t (y|k)$

Dal lemma 13.12 sappiamo che $1 - (\varepsilon|k)^t (z|k) \neq 0$, per cui possiamo scrivere

$$\lambda = \frac{\gamma_{k+1} - (\varepsilon|k)^t (y|k)}{1 - (\varepsilon|k)^t (z|k)}$$

Da ciò troviamo $v = (y|k) - \lambda(z|k)$.

Nota 13.14. Sia dato un vettore $b \in \mathbb{R}^n$ per il quale poniamo

$$(b|k) := \begin{pmatrix} b^1 \\ \vdots \\ b^k \end{pmatrix}$$

ed $(x|k) := (\Omega|k)^{-1}(b|k)$.

Possiamo determinare $(x|k)$ con essenzialmente lo stesso algoritmo della nota 13.13, in cui però usiamo la conoscenza di $(y|k)$:

(1) Ovviamente $(x|1) = (b|1) = b^1$.

(2) Assumiamo di conoscere $(x|k)$ e tramite l'algoritmo della nota 13.13 anche $(y|k)$ e $(z|k)$. $(x|k+1)$ è la soluzione dell'equazione $(\Omega|k+1)(x|k+1) = (b|k+1)$ che possiamo scrivere nella forma:

$$\begin{pmatrix} (\Omega|k) & (\varepsilon|k) \\ (\varepsilon|k)^t & 1 \end{pmatrix} \begin{pmatrix} w \\ \mu \end{pmatrix} = \begin{pmatrix} (b|k) \\ b^{k+1} \end{pmatrix}$$

con w e μ da determinare. Ciò equivale a

$$(\Omega|k)w + \mu(\varepsilon|k) = (b|k)$$

$$(\varepsilon|k)^t w + \mu = b^{k+1}$$

ovvero

$$w = (\Omega|k)^{-1}(b|k) - \mu(\Omega|k)^{-1}(\varepsilon|k) = (x|k) - \mu(z|k)$$

$$\mu = b^{k+1} - (\varepsilon|k)^t w$$

da cui troviamo

$$\mu = b^{k+1} - (\varepsilon|k)^t (x|k) + \mu(\varepsilon|k)^t (z|k)$$

cosicché

$$\mu = (1 - (\varepsilon|k)^t(z|k)) = b^{k+1} - (\varepsilon|k)^t(x|k)$$

Come nella nota 13.13 osserviamo che $1 - (\varepsilon|k)^t(z|k) \neq 0$, per cui possiamo scrivere

$$\mu = \frac{b_{k+1} - (\varepsilon|k)^t(x|k)}{1 - (\varepsilon|k)^t(z|k)}$$

$$w = (x|k) - \mu(z|k)$$

Nota 13.15. Combinando le note 13.13 e 13.14 possiamo tradurre l'algoritmo della nota 13.13 in un programma in Matlab:

```
function x=toeplitz(gamma,b)

    if gamma(1)~=1, fprintf('gamma_0 deve essere 1\n'); return, end

    k=length(b); gamma=gamma(2:k);
    x=[]; y=[]; x(1)=b(1); y(1)=gamma(1);

    for j=1:1:(k-1)

        epsilonj=flipud(gamma(1:j)');
        zetaj=flipud(y(1:j));
        mu=(b(j+1)-epsilonj'*x(1:j))/(1-epsilonj'*zetaj);
        w=x(1:j)-mu*zetaj;
        x(1:j+1,1)=[w ; mu];

        if j<(k-1)
            lambda=(gamma(j+1)-epsilonj'*y(1:j))/(1-epsilonj'*zetaj);
            v=y(1:j)-lambda*zetaj;
            y(1:j+1,1)=[v ; lambda];
        end

    end

end
```

Osservazione 13.16. Dalla nota 13.14 otteniamo in particolare un algoritmo ricorsivo per il calcolo di $(\Omega|n)^{-1}$. Un algoritmo più veloce, noto come algoritmo di Trench, è presentato ad esempio in Hanke-Bourgeois, pagg. 70-72, oppure in Golub/Van Loan, pagg. 195-199.

Esempio 13.17. Siano $\gamma_0 = 1, \gamma_1 = 4, \gamma_2 = 3, \gamma_3 = 2$ e $b = \begin{pmatrix} 7 \\ 2 \\ 3 \\ 1 \end{pmatrix}$.

Calcoliamo $(x|4) = (\Omega|4)^{-1} b$.

(1) Inizialmente abbiamo

$$(y|1) = \gamma_1 = 4; (\varepsilon|1) = 4; (z|1) = 4 \text{ e quindi } (x|1) = 4.$$

(2) Per $k = 1$ otteniamo

$$\mu = \frac{2 - 4 \cdot 7}{1 - 4 \cdot 4} = \frac{26}{15} \text{ e } w = 7 - 4 \cdot \frac{26}{15} = \frac{1}{15}$$

per cui $(x|2) = \begin{pmatrix} w \\ \mu \end{pmatrix} = \begin{pmatrix} 1/15 \\ 26/15 \end{pmatrix}$

$$\lambda = \frac{3 - 4 \cdot 4}{1 - 4 \cdot 4} = \frac{13}{15} \text{ e } v = 4 - \frac{13}{15} \cdot 4 = \frac{8}{15}$$

per cui $(y|2) = \begin{pmatrix} 8/15 \\ 13/15 \end{pmatrix}$

(3) Per $k = 2$ otteniamo

$$(\varepsilon|2) = \begin{pmatrix} 3 \\ 4 \end{pmatrix}; (z|2) = \begin{pmatrix} 13/15 \\ 8/15 \end{pmatrix}$$

$$\mu = \frac{3 - \frac{1}{15} \left[(3 \ 4) \cdot \begin{pmatrix} 1 \\ 26 \end{pmatrix} \right]}{1 - \frac{1}{15} \left[(3 \ 4) \cdot \begin{pmatrix} 13 \\ 8 \end{pmatrix} \right]} = \frac{31}{28}$$

$$w = \frac{1}{15} \left[\begin{pmatrix} 1 \\ 26 \end{pmatrix} - \frac{31}{28} \begin{pmatrix} 13 \\ 8 \end{pmatrix} \right] = \begin{pmatrix} -25/28 \\ 8/7 \end{pmatrix}$$

per cui $(x|3) = \begin{pmatrix} w \\ \mu \end{pmatrix} = \begin{pmatrix} -25/28 \\ 8/7 \\ 31/28 \end{pmatrix}$

$$\lambda = \frac{2 - \frac{1}{15} \left[(3 \ 4) \cdot \begin{pmatrix} 8 \\ 13 \end{pmatrix} \right]}{1 - \frac{1}{15} \left[(3 \ 4) \cdot \begin{pmatrix} 13 \\ 8 \end{pmatrix} \right]} = \frac{23}{28}$$

$$v = \frac{1}{15} \left[\begin{pmatrix} 8 \\ 13 \end{pmatrix} - \frac{23}{28} \begin{pmatrix} 13 \\ 8 \end{pmatrix} \right] = \begin{pmatrix} -5/28 \\ 3/7 \end{pmatrix}$$

da cui otteniamo $(y|3) = \begin{pmatrix} v \\ \lambda \end{pmatrix} = \begin{pmatrix} -5/28 \\ 3/7 \\ 23/28 \end{pmatrix}$

(4) Per $k = 3$ abbiamo infine

$$(\varepsilon|3) = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}; (z|3) = \begin{pmatrix} 23/28 \\ 3/7 \\ -5/28 \end{pmatrix}$$

$$\mu = \frac{1 - \frac{1}{28} \left[(2 \ 3 \ 4) \cdot \begin{pmatrix} -25 \\ 32 \\ 31 \end{pmatrix} \right]}{1 - \frac{1}{28} \left[(2 \ 3 \ 4) \cdot \begin{pmatrix} 23 \\ 12 \\ -5 \end{pmatrix} \right]} = \frac{71}{17}$$

$$w = \frac{1}{28} \left[\begin{pmatrix} -25 \\ 32 \\ 31 \end{pmatrix} - \frac{71}{17} \begin{pmatrix} 23 \\ 12 \\ -5 \end{pmatrix} \right] = \begin{pmatrix} -147/34 \\ -11/17 \\ 63/34 \end{pmatrix}$$

e la soluzione è $(x|4) = \begin{pmatrix} w \\ \lambda \end{pmatrix} = \begin{pmatrix} -147/34 \\ -11/17 \\ 63/34 \\ 71/17 \end{pmatrix}$

Bibliografia

- P. Billingsley:** Convergence of probability measures. Willey 1999.
- P. Brockwell/R. Davis:** Time series - theory and methods. Springer 2006.
- J. Elstrodt:** Maß- und Integrationstheorie. Springer 2005.
- R. Engelking:** General topology. Heldermann 1989.
- G. Golub/C. Van Loan:** Matrix computations. Johns Hopkins UP 1996.
- M. Hanke-Bourgeois:** Grundlagen der numerischen Mathematik und des wissenschaftlichen Rechnens. Teubner 2006.
- H. Heuser:** Funktionalanalysis. Teubner 1986.
- F. Hirzebruch/W. Scharlau:** Einführung in die Funktionalanalysis. Bibl. Inst. 1971.
- Y. Katznelson:** An introduction to harmonic analysis. Dover 1976.
- J. Kreiß/G. Neuhaus:** Einführung in die Zeitreihenanalyse. Springer 2006.
- K. Neusser:** Zeitreihenanalyse in den Wirtschaftswissenschaften. Teubner 2006.
- M. Priestley:** Spectral analysis and time series. Academic Press 1989.
- R. Schlittgen/B. Streitberg:** Zeitreihenanalyse. Oldenbourg 1989.
- H. Schröder:** Funktionalanalysis. Deutsch 2000.
- J. Weidmann:** Lineare Operatoren in Hilberträumen. Teubner 1976.