Rumus eksplisit untuk sudut antara dua sub-ruang dari suatu ruang hasilkali dalam[1]

Hendra Gunawan[2]

Departemen Matematika
Institut Teknologi Bandung
Bandung 40132, Indonesia

[1] Dipresentasikan pada Konferensi Nasional Matematika XII, Bali, 23-27 Juli 2004.

[2] Bekerjasama dengan O. Neswan dan W. Setya-Budhi

Abstrak

Konsep sudut antara dua subruang dari ruang Euclid $\mathbf{R}^d$ telah menarik perhatian matematikawan sejak tahun 1950-an [3]. Dalam statistika, misalnya, dikenal sudut kanonis yang merupakan ukuran ketergantungan suatu himpunan peubah acak pada himpunan peubah acak lainnya [1]. Beberapa hasil penelitian terakhir tentang sudut antara dua subruang dan topik yang terkait dapat ditemui misalnya di [4, 8, 12, 13, 14]. Khususnya, di [13], Risteksi dan Trenčevski menawarkan suatu definisi sudut antara dua subruang dari $\mathbf{R}^d$ yang lebih bersifat geometris dan menjelaskan kaitannya dengan sudut kanonis. Definisi mereka, sayangnya, dirumuskan berdasarkan pada suatu perumuman ketaksamaan Cauchy-Schwarz yang salah. Dalam makalah ini, definisi mereka akan ditinjau ulang dan diperbaiki. Pada saat yang sama, ruang semestanya akan diperluas menjadi ruang hasilkali dalam sebarang. Suatu perumuman ketaksamaan Cauchy-Schwarz yang merupakan perbaikan dari ketaksamaan yang ada di [13] juga diperoleh.

Abstract

The notion of angles between two subspaces of the Euclidean space $\mathbf{R}^d$ has been studied by many researchers since the 1950's or even earlier (see [3]. In statistics, canonical (or principal) angles are studied as measures of dependency of one set of random variables on another (see [1]). Some recent works on angles between subspaces and related topics can be found in, for examples, [4, 8, 12, 13, 14]. Particularly, in [13], Risteski and Trenčevski introduced a more geometrical definition of angle between two subspaces of $\mathbf{R}^d$ and explained its connection with canonical angles. Their definition of the angle, however, is based on a generalized Cauchy-Schwarz inequality which we found incorrect. In this paper, their definition will be reviewed and refined. At the same time, the ambient space is extended to any real inner product space. A generalization of the Cauchy-Schwarz inequality that serves as a correction for the one in [13] is also obtained.

Pendahuluan

Misalkan $(X, \langle\cdot,\cdot\rangle)$ suatu ruang hasilkali dalam real. Diberikan dua subruang U dan V dari X, yang tak kosong dan berdimensi hingga, dengan $\dim(U) \leq \dim(V)$, kita ingin mendefinisikan sudut θ antara U dan V.

Definisi sudut tersebut harus dapat dipandang sebagai perumuman dari definisi sudut yang 'biasa' antara

(a) subruang berdimensi 1 dan subruang berdimensi q, dan

(b) dua subruang berdimensi p yang beririsan pada suatu subruang berdimensi $(p-1)$ yang sama.

Di sini kita asumsikan bahwa $\theta \in [0, \frac{\pi}{2}]$.

Dengan memisalkan $U = \text{span}\{u_1, \ldots, u_p\}$ dan $V = \text{span}\{v_1, \ldots, v_q\}$ dengan $p \leq q$, Risteski dan Trenčevski mendefinisikan sudut θ antara U dan V melalui

$$\cos^2\theta := \frac{\det(MM^{\mathsf{T}})}{\det[\langle u_i, u_j\rangle].\det[\langle v_k, v_l\rangle]}. \qquad (1)$$

Rumus ini didasarkan pada 'ketaksamaan' Cauchy-Schwarz yang diperumum

$$\det(MM^{\mathsf{T}}) \leq \det[\langle u_i, u_j\rangle].\det[\langle v_k, v_l\rangle], \qquad (2)$$

dengan $M := [\langle u_i, v_k\rangle]$.

Namun, 'ketaksamaan' ini benar hanya dalam kasus

(a) $p = q$, di mana (2) tereduksi menjadi ketaksamaan Cauchy-Schwarz diperumum versi Kurepa [9], dan

(b) $\{v_1, \ldots, v_q\}$ ortonormal.

Secara umum, rumus (1) tidak masuk akal karena bentuk di ruas kanan tidak selalu merupakan bilangan di $[0, 1]$.

Untuk memperlihatkan bahwa (2) salah secara umum, ambil sebagai contoh $X = \mathbf{R}^3$ (dengan hasilkali dalam biasa), $U = \text{span}\{u\}$ dengan $u = (1, 0, 0)$, dan $V = \text{span}\{v_1, v_2\}$ dengan $v_1 = (\frac{1}{2}, \frac{1}{2}, 0)$ dan $v_2 = (\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$.

Menurut (2), kita mempunyai 'ketaksamaan'

$$\langle u, v_1 \rangle^2 + \langle u, v_2 \rangle^2 \leq \|u\|^2 \det[\langle v_i, v_j \rangle]$$

yang setara dengan

$$\frac{1}{2} \leq \frac{3}{8}.$$

Contoh ini memperlihatkan bahwa (2) salah sekalipun dalam kasus $\{u_1, \ldots, u_p\}$ ortonormal dan $\{v_1, \ldots, v_q\}$ ortogonal.

Revisi Rumus Risteski dan Trenčevski

Untuk merivisi rumus (1), kita perlu mengingat kembali bahwa definisi sudut yang kita inginkan harus merupakan perumuman definisi sudut yang 'biasa' antara

(a) subruang berdimensi 1 dan subruang berdimensi q, dan

(b) dua subruang berdimensi p yang beririsan pada suatu subruang berdimensi $(p-1)$ yang sama.

Oleh karena itu, marilah kita tinjau bagaimana sudut antara dua subruang dalam kasus (a) dan (b) biasanya didefinisikan.

Dalam kasus (a), jika $U = \text{span}\{u\}$ subruang berdimensi 1 dan $V = \text{span}\{v_1, \ldots, v_q\}$ subruang berdimensi q, maka sudut θ antara U dan V didefinisikan melalui

$$\cos^2 \theta := \frac{\langle u, u_V \rangle^2}{\|u\|^2 \|u_V\|^2} \tag{3}$$

dengan u_V menyatakan proyeksi (ortogonal) dari u pada V dan $\|\cdot\| := \langle \cdot, \cdot \rangle^{\frac{1}{2}}$ menyatakan norm pada X.

Dalam kasus (b), jika $U = \text{span}\{u, w_2, \ldots, w_p\}$ dan $V = \text{span}\{v, w_2, \ldots, w_p\}$ dua subruang berdimensi p ($p \geq 2$) yang beririsan pada subruang $W = \text{span}\{w_2, \ldots, w_p\}$ yang berdimensi $p - 1$, maka sudut θ antara U dan V didefinisikan melalui

$$\cos^2 \theta := \frac{\langle u_W^\perp, v_W^\perp \rangle^2}{\|u_W^\perp\|^2 \|v_W^\perp\|^2} \tag{4}$$

dengan $u_W^\perp$ dan $v_W^\perp$ menyatakan komplemen ortogonal dari u dan v terhadap W.

Perhatikan bahwa dalam (a), kita dapat menuliskan $u = u_V + u_V^{\perp}$ dengan $u_V^{\perp}$ menyatakan komplemen ortogonal dari u terhadap V, sehingga (3) menjadi

$$\cos^2\theta = \frac{\|u_V\|^2}{\|u\|^2}.$$

Ini mengatakan bahwa nilai $\cos\theta$ sama dengan rasio antara panjang proyeksi u pada V dan panjang u sendiri.

Serupa dengan itu, dalam (b), kita dapat menunjukkan bahwa $\cos\theta$ sama dengan rasio antara volume paralelpipedium berdimensi p yang direntang oleh proyeksi dari $u, w_2, \ldots, w_p$ pada V dan volume paralelpipedium berdimensi p yang direntang oleh $u, w_2, \ldots, w_p$.

Menggunakan p-norm baku pada X, sudut θ antara subruang $U = \text{span}\{u_1, \ldots, u_p\}$ dan $V = \text{span}\{v_1, \ldots, v_q\}$ (dengan $p \leq q$) kita definisikan melalui

$$\cos^2 \theta := \frac{\|\text{proj}_V u_1, \ldots, \text{proj}_V u_p\|^2}{\|u_1, \ldots, u_p\|^2}, \qquad (5)$$

dengan $\|\cdot, \ldots, \cdot\|$ menyatakan p-norm baku pada X. (Secara geometris, $\|u_1, \ldots, u_p\|$ menyatakan volume paralelpipedium berdimensi p yang direntang oleh $u_1, \ldots, u_p$; lihat apendiks).

Fakta. *Rasio di ruas kanan* (5) *merupakan bilangan di* $[0, 1]$ *dan tak bergantung pada pemilihan basis untuk U dan V.*

Bukti. Pertama perhatikan bahwa proyeksi u_i pada V tak bergantung pada pemilihan basis untuk V. Lalu, karena proyeksi bersifat linear, rasio (5) tidak berubah jika kita

(a) menukar u_i dan u_j,

(b) mengganti u_i dengan $u_i + \alpha u_j$, atau

(c) mengganti u_i dengan αu_i with $\alpha \neq 0$.

Karena itu rasio (5) invarian terhadap perubahan basis untuk U.

Selanjutnya, dengan asumsi bahwa $\{u_1, \ldots, u_p\}$ ortonormal, kita punyai $\|u_1, \ldots, u_p\| = 1$ dan $\|\text{proj}_V u_1, \ldots, \text{proj}_V u_p\| \leq 1$ sebab $\|\text{proj}_V u_i\| \leq \|u_i\| = 1$ untuk tiap $i = 1, \ldots, p$. Jadi, rasio (5) merupakan suatu bilangan di $[0, 1]$.

Rumus Eksplisit untuk $\cos\theta$

Dari (5), kita dapat memperoleh rumus eksplisit untuk $\cos\theta$ dinyatakan dalam $u_1, \ldots, u_p$ and $v_1, \ldots, v_q$. Khususnya, jika $\{u_1, \ldots, u_p\}$ dan $\{v_1, \ldots, v_q\}$ ortonormal, maka (5) menjadi

$$\cos^2\theta = \det(MM^\mathsf{T}). \tag{6}$$

Jika $\{v_1, \ldots, v_q\}$ saja yang ortonormal, maka (5) menjadi

$$\cos^2\theta = \frac{\det(MM^\mathsf{T})}{\det[\langle u_i, u_j\rangle]}. \tag{7}$$

Perhatikan bahwa dalam hal $p = q$, kita punya

$$\det(MM^\mathsf{T}) = \det M. \det M^\mathsf{T} = \det{}^2 M,$$

sehingga (6) dapat disederhanakan menjadi

$$\cos\theta = |\det M|. \tag{8}$$

Catatan. Di sini $M := [\langle u_i, v_k\rangle]$ merupakan matriks $p \times q$.

Secara umum, kita mempunyai rumus

$$\cos^2\theta = \frac{\det(M\tilde{M}^{\mathsf{T}})}{\det[\langle u_i, u_j\rangle].\det^p[\langle v_k, v_l\rangle]}, \qquad (9)$$

dengan $M = [\langle u_i, v_k\rangle]$ dan

$$\tilde{M} := [\langle u_i, v_k | v_{i_2(k)}, \ldots, v_{i_q(k)}] \qquad (10)$$

di mana $\{i_2(k), \ldots, i_q(k)\} = \{1, 2, \ldots, q\} \setminus \{k\}$, $k = 1, 2, \ldots, q$.

Di sini $\langle \cdot, \cdot | \cdot, \ldots, \cdot\rangle$ menyatakan hasilkali dalam-q baku (lihat apendiks).

Rumus (9) merupakan revisi dari Rumus (1). Perhatikan bahwa dalam hal $\{v_1, \ldots, v_q\}$ ortonormal, kita peroleh Rumus (7).

Sebagai akibat dari Rumus (9) dan Fakta sebelumnya, kita peroleh ketaksamaan berikut yang merupakan perumuman dari Ketaksamaan Cauchy-Schwarz, dan merupakan revisi dari Ketaksamaan (2).

Teorema. *Untuk dua himpunan bebas linear sebarang* $\{u_1, \ldots, u_p\}$ *dan* $\{v_1, \ldots, v_q\}$ *di* X *dengan* $p \leq q$ *berlaku*

$$\mathrm{det}(M\tilde{M}^{\mathsf{T}}) \leq \mathrm{det}[\langle u_i, u_j\rangle].\mathrm{det}^p[\langle v_k, v_l\rangle],$$

di mana $M = [\langle u_i, v_k\rangle]$ *dan* $\tilde{M}$ *sebagaimana diberikan dalam Rumus* (10)*. Lebih jauh, kesamaan berlaku jika dan hanya jika subruang yang direntang oleh* $\{u_1, \ldots, u_p\}$ *termuat dalam subruang yang direntang oleh* $\{v_1, \ldots, v_q\}$*.*

Apendiks

Pada X, hasilkali dalam-n baku $\langle \cdot, \cdot | \cdot, \ldots, \cdot \rangle$ didefinisikan sebagai

$$\langle x_0, x_1 | x_2, \ldots, x_n \rangle := \begin{vmatrix} \langle x_0, x_1 \rangle & \langle x_0, x_2 \rangle & \ldots & \langle x_0, x_n \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \ldots & \langle x_2, x_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1 \rangle & \langle x_n, x_2 \rangle & \ldots & \langle x_n, x_n \rangle \end{vmatrix},$$

dan norm-n baku $\|\cdot, \ldots, \cdot\|$ didefinisikan sebagai

$$\|x_1, x_2, \ldots, x_n\| := \langle x_1, x_1 | x_2, \ldots, x_n \rangle^{\frac{1}{2}}$$

(lihat [6] atau [10]). Jika $n = 1$, hasilkali dalam-1 baku dipahami sebagai hasilkali dalam, sementara norm-1 baku merupakan norm pada X.

Catat bahwa $\langle x_1, x_1 | x_2, \ldots, x_n \rangle = \det[\langle x_i, x_j \rangle]$ tak lain merupakan determinan Gram yang dibangun oleh x_1, x_2, ..., x_n (lihat [5] dan [11]). Secara geometris, $\|x_1, \ldots, x_n\|$ menyatakan volume paralelpipedium berdimensi n yang direntang oleh $x_1, \ldots, x_n$.

Diberikan $u \in X$ dan $V = \text{span}\{v_1, \ldots, v_q\}$ subruang dari X, proyeksi dari u pada V dapat dinyatakan sebagai

$$\text{proj}_V u = \sum_{k=1}^{q} \frac{\langle u, v_k | v_{i_2(k)}, \ldots, v_{i_q(k)} \rangle}{\|v_1, v_2, \ldots, v_q\|^2} v_k$$

dengan $\{i_2(k), \ldots, i_q(k)\} = \{1, \ldots, q\} \setminus \{k\},\ k = 1, \ldots, q$.

Referensi

[1] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, Inc., New York, 1958.

[2] A.L. Brown and A. Page, *Elements of Functional Analysis*, Van Nostrand Reinhold Co., London, 1970.

[3] C. Davies and W. Kahan, "The rotation of eigenvectors by a perturbation. III", *SIAM J. Numer. Anal.* **7** (1970), 1–46.

[4] S. Fedorov, "Angle between subspaces of analytic and antianalytic functions in weighted L_2 space on a boundary of a multiply connected domain" in *Operator Theory, System Theory and Related Topics*, Beer-Sheva/Rehovot, 1997, 229–256.

[5] F.R. Gantmacher, *The Theory of Matrices*, Vol. I, Chelsea Publ. Co., New York (1960), 247–256.

[6] H. Gunawan, "On n-inner products, n-norms, and the Cauchy-Schwarz inequality", *Sci. Math. Jpn.* **5** (2001), 47–54.

[7] H. Gunawan, "The space of p-summable sequences and its natural n-norm", *Bull. Austral. Math. Soc.* **64** (2001), 137–147.

[8] A.V. Knyazev and M.E. Argentati, “Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates”, *SIAM J. Sci. Comput.* **23** (2002), 2008–2040.

[9] S. Kurepa, “On the Buniakowsky-Cauchy-Schwarz inequality”, *Glasnik Mat. Ser. III* **1**(21) (1966), 147–158.

[10] A. Misiak, “n-inner product spaces”, *Math. Nachr.* **140** (1989), 299–319.

[11] D.S. Mitrinović, J.E. Pečarić and A.M. Fink, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht (1993), 595–603.

[12] V. Rakočević and H.K. Wimmer, “A variational characterization of canonical angles between subspaces”, Preprint.

[13] I.B. Risteski and K.G. Trenčevski, “Principal values and principal subspaces of two subspaces of vector spaces with inner product”, *Beiträge Algebra Geom.* **42** (2001), 289–300.

[14] H.K. Wimmer, “Canonical angles of unitary spaces and perturbations of direct complements”, *Linear Algebra Appl.* **287** (1999), 373–379.