Topics in Computer Science Spring 2025
COS 470 Extra Scribe 1

Objectives

- Objective 1
- Objective 2
- Objective 3

---

1 Theorem

if $u_1, u_2, ..., u_u$ Orthonormal and $v = \sum c_i u_i$ then $c_i =< u_i, v >$

1.1 Proof

$$\begin{aligned}
< u_i, v > &=< u_i, \sum_{j=1}^{n} c_j u_j \\
&= \sum c_j < u_{ji} u_i > \\
&= \sum_{j=1}^{n} c_j f_{ij} \\
&= c_i
\end{aligned}$$

Based on: $< ax + y, z >= a < x, z > + < y, z >$
Corollary: if $u = \sum a_i u_i \wedge \vee = \sum b_i u_i$

1.2 Proof

$$
\begin{aligned}
< u, v > &=< a_i u_i, b_i u_i > \\
&= a_1 < u_1, \sum b_i u_i > + a_2 < u_2, v > + ... \\
&= a_1 b_1 < u_1, u_1 > + a_1 b_2 < u_1, u_2 > + ... \\
&+ a_2 b1 < u_2, u_1 > + ... \\
< u, v > &=< \sum a_i u_i, v > \\
&= \sum a_i < u_i, v > \\
&= \sum a_i < v_1 u_i > \\
&= \sum a_i b_i
\end{aligned}
$$

1.3 Definitions

1.3.1 Def 1

$||v|| = \sqrt{< v, v >} = \sqrt{v^t v}$

1.3.2 Def 2

$O \in \mathbb{R}^n * n$ is or orthogonal if Qi from orthogonal set.

1.3.3 Def 3

Orthogonal set $\{u_1, u_2, ..., u_u\}$
Orthogonal unit vectors

$$
u_i^t u_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
$$

1.3.4 Def 4

Gwen avg $\{v_1, v_2, ..., v_u\}$ orthogonal
$u_i = \frac{v_i}{||v_i||}$ are normalized, hence orthogonal
Ex:

$$
\begin{aligned}
v_1 &= (1, 1, 1)^\top \rightarrow \quad u_1 = \frac{1}{\sqrt{3}} v_1 \\
v_2 &= (2, 1, -3)^\top \rightarrow \quad u_2 = \frac{1}{\sqrt{14}} v_2 \\
v_3 &= (4, -5, 1)^\top \rightarrow \quad u_3 = \frac{1}{\sqrt{42}} v_3
\end{aligned}
$$

1.3.5 Def 5

$a|b$ iff $\exists k \in \mathbb{Z} : b = ak$

- b is a multiple of a
- a is a divisor of b
- a is a factor of b

1.3.6 Def 6

$n > 1$ is prime is only divisors a $1 \wedge n$. (no factors)

1.3.7 Def 7

Composite is not prime (has at least 1 factor).

1.3.8 Def 8

If a/n $\wedge a/m$, then a is a common factor.

1.3.9 Def 9

$\gcd(m, n) = d \Leftrightarrow d|m \wedge d|u$ and if $h|m \wedge h|n$, then $d \leqslant h$

2 Theorem

if $h|n \wedge h|m$, then $\forall i, jh|(in + jm)$.

2.1 Proof

$n = nk, \wedge m = hk_2$ then

$$\begin{aligned} in + jm &= ihk_1 + jhk_2 \\ &= h(ik_1 + jk_2) \\ \rightarrow & h|(in + jm) \end{aligned}$$

2.2 Definition

(Quotient) $\forall m, n : m \neq 0$ $q = \lfloor \frac{n}{m} \rfloor$ and (remainder) of dividing n by m is $r = n - qm$. ($r = n \bmod m$)
i.e.

$$\boxed{n = mq + r \quad \text{such that} \quad \begin{cases} 0 & \text{if } 0 \leqslant r < m \quad m > 0 \\ m & \text{if } m \leqslant r \leqslant 0 \quad m < 0 \end{cases}}$$

3 Theorem

let $m, n \in \mathbb{Z} : m \neq 0 V n \neq 0$ and (not both 0) let $d = min\{in + jm \text{such that} i, j \in \mathbb{Z} \wedge in + jm > 0\}$ then $d = gcd(m, n)$.

3.1 Proof

$d = in + jm$.(smallest linear combo of $m \wedge n$, moreover, $n = qd + r$ with $0 \leqslant r < d$. Therefor we have

$$\begin{aligned} n &= q(in + jm) + r \\ &= qin + qjm + r \\ r &= n(1 - qi) + m(qi) \end{aligned}$$

Since r is linear combo but d is minimal $\wedge r < d$, then $r = 0$, $\rightarrow d|n$. Sumilarly $d|m \rightarrow d$ is common divisor $\rightarrow d \leqslant gcd(m, n)$. Also since $gcd(m, n)$ divides $m \wedge n$ and $d = in + jm$, then $gcd(m, n)|d \rightarrow d = gcd(m, n)k \rightarrow d \geqslant acd(m, n)$ **Corollary:** if $h|n \wedge h|m$, then $h|gcd(m, n)$

3.2 Proof

$gcd(m, n) = am + bn$
$\therefore$ since $h|m \wedge h|n$ then $m = hk_1 \wedge n = hk_2$ then

$$\begin{aligned} gcd(m, n) &= hak_1 + hbk_2 \\ &= h(ak_1 + dk_2) \end{aligned}$$

4 Theorem

Suppose $n \geqslant 0 \quad m > 0$. Let $r = n \quad modm$ then $gcd(m, n) = gcd(m, r)$.

4.1 Proof

show $gcd(m, n)|gcd(m, n)$ and vice versa. Let $d_1 = gcd(m, n)$, then $d_1|m \wedge d_1/n \rightarrow m = d_1k_1 \wedge n = d_1k_2$. Also $n = mq + r \quad \exists q_1 r{:}0 \leqslant r < m$

5 Gradient Descent

- Compute sequence of vectors in $\mathbb{R}^s$ aiming to converge to a vector that minimizes cost
- if $\Delta p \approx 0$,then Taylor Series gives $Cost(p + \Delta p) \approx Cost(p) + \sum_{r=1}^{s} \frac{\partial Cost(p)}{\partial pr}.\Delta pr$. $Cost(p) + \frac{\partial C(p)}{\partial p} + \frac{\partial C(p)}{\partial p_2} + ... + \frac{\partial C(p)}{\partial p_2}$
- Denote $\nabla C(p) \in \mathbb{R}^s$, gradient (vector of partial derivatives).
- $\frac{\partial C(p)}{\partial pi}$ = how much Cost changes in the "direction" pi.

- $(\nabla C(p))_r = \frac{\partial C(p)}{\partial p_r}$
- $Cost(p + \Delta p) \approx Cost(p) + \nabla Cost(p)^t \Delta p$.

$$\boxed{Cost(p + \Delta p) \approx C(p) + \nabla C(p)^t \Delta p}$$

- Choose Δp to make $\nabla C(p)^t \Delta p$ as negative as possible.
- Cauchy - Schwarz Inequality.
  $|x^t y| \leqslant ||x||_2 * ||y||_2$
  $\therefore -||x||_2 * ||y||_2 \leqslant x^t y \leqslant ||x||_2 * ||y||_2$
  $-||x||_2 * ||y||_2$ most negtive $x^t y$ can be.
  $\rightarrow$ When $x = -y$.

Note: $|x^y| = |x * y| = ||x||_2 * ||y||_2 * cos0 \quad x = \lambda y$, then $|\lambda y * y| = |\lambda| * ||y||^2 = ||y|| * ||y|| * |cos0|$
$||x|| = \sqrt{x^t x}$
$||y|| =$