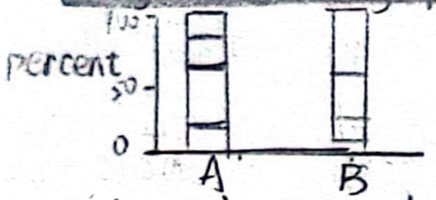
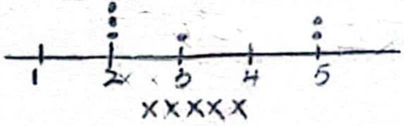
AP统计

1 **segment bar graph** **dotplots**



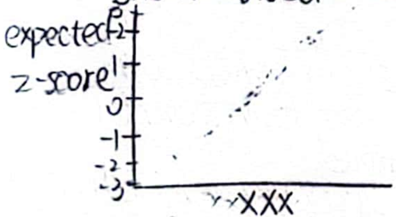


2. **describe overall pattern**: shape/center/spread
describe striking departure: outlier

3. **unimodal**: single peak
bimodal double peaks
multimodal: more than 2 peaks

6 **density curve** doesn't reflect outliers, and it is an approximation

7 • **for normal distribution**
≈68% of data fall within σ
≈95% of data fall within 2σ
≈99.7% of data fall within 3σ
• if the points on a **normal probability plot** close to a straight line, the data are approximately normally distributed.



if a is β percentile
$\varphi(c)=\beta$
we plot (a, c)

outliers appear as points that are far away from overall pattern of the plot

11. • **regression line** 回归曲线: $\hat{y}=ax+b$
描述当自变量变化时因变量的变化程度
• **extrapolation**: 外推法
用interval之外的自变量推因变量,不一定 accurate
• **residual**: 推导值与真实值之差 $y-\hat{y}$
• **least square regression line**:
make the sum of squared residuals as small as possible
• **residual plot**:
a scatterplot of residuals against explanatory variable
when obvious pattern occured in residual plot, model isn't appropriate

4. **pth percentile**: p% 数据比其小.

5. **standardized score (z-score)**

$$z=\frac{x-\bar{x}}{\sigma}$$

one data is z standard deviation below/above the mean.
用这个比较两个的相对位置.

8. **explanatory variable** "自变量"
response variable "因变量"

9. **scatter plot**

explanatory variable: horizontal axis,
response variable: vertical axis,
every data as a point

描述: direction, form, strength
从哪个方向到哪个方向 / 直线曲线 clusters outliers / 相关性

10. **相关系数**: $r=\frac{1}{n-1}\sum z_x z_y$
• doesn't imply causation
• ignores the distinction between explanatory & response variables
• isn't affected by changes in unit of measurement
• isn't resistant

• **standard deviation of residuals**

$$s=\sqrt{\frac{\Sigma(y_i-\hat{y}_i)^2}{n-2}}$$ ("误差")

• **coefficient of determination**

$$r^2=1-\frac{\Sigma \text{residuals}^2}{\Sigma(y_i-\bar{y})^2}$$

如果用 y 估计，我们能得到一个 corresponding squared residuals. 这个数据与最小值的比反映了有多少程度的因变量是不被自变量影响的

- for X, Y, least-squared regression line is $\hat{y}=ax+b$ where $a=r\frac{s_y}{s_x}$, $b=\bar{y}-b\bar{x}$

13. experiment (VS observational study)

- when goal is to understand cause and effect, experiments are only source of fully convincing data
- confounding: two variables are associated in such a way that their effects on a response variable cannot be distinguished from each other
- treatment: specific condition applied to individuals in an experiment

  experimental unit: smallest collection of individuals to which treatments are applied

  subject: when units are human beings

  factor: explanatory variables

  level: treatment formed by combining a specific value
- priciples of experimental design

  comparison + random assignment + control + replication
- completely randomized design

  the experimental units are assigned to the treatment completely by chance (not necessarily equal numbers)
- control group: 对照组
- double-blind experiment: neither the subjects nor those who interact with them and measure the response variable know which treatment a subject received
- statistically significant: an observed effect so large that it would rarely occur by chance

12 数据采集

- convenience survey 便利抽样

  快速得结论但 bias
- stratified random sample 分层抽样

  strata: 层

  starting by classify the population in to groups of similar individuals (strata), the choose a seperate SRS in each stratum and combine these SRSs to form the sample

  it works best when the individuals within each stratum are similar with respect to what is being measured and when there are large differences between strata
- cluster sample 聚类样本

  start by classifying the population into groups of individuals, that are located near to each other, (cluster), then choose an SRS of the clusters
- difference between stratified random & cluster sample

  前者分成多个小各目，后者分为多个小 population
- undercoverage occurs when some members of the population cannot be chosen in a sample
- nonrespone occurs when an individual chosen for the sample can't be contacted or refuses to participate

  nonrespone ≠ voluntary response

  前者已选好 sample，后者在选 sample
- response bias

  a systematic pattern of inaccurrate answers in a survey

  wording of question

- block: a group of experimental units that are known before the experiment to be similar in some way that is expected to affect the response to the treatments

  randomized block design

  the random assignment

experimental units to treatment is carried out separately within each block

matched pairs design: a common form of blocking for comparing just two treatments

15. sample space S: set of all possible outcomes of a process.

probability model: sample space S + probabilities

16. - parameter describes population. statistic describes sample

in population μ σ p N

in sample: $\bar{x}$ s $\hat{p}$ n

- sample distribution: the distribution of values taken by statistic in all possible samples of the same size from the same population
- variability of a statistic: spread of its sampling distribution, determined by sizes of samples
- sample distribution of $\hat{p}$

$\mu_{\hat{p}} = p$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ for $n \leq \frac{1}{10}N$

if $np > 10$, $nq > 10$, approximately normal

sample distribution of $\bar{x}$

$\mu_{\bar{x}} = \mu$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ for $n \leq \frac{1}{10}N$

if $n \geq 30$, 中心极限定理CLT distribution will be approximately normal

- CI for μ: $\bar{x} \pm t_{critical} \frac{s}{\sqrt{n}}$

$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

注意，t的自由度是$n-1$，如果题目要求自由度没有，用the greatest df available that is less than your desired

conditions

① well-design random sample or

14. make conclusion

-

	individuals randomly assigned to groups: Y	N
were individuals randomly selected: Y	P C&E	P
N	C&E	

P: inference about population

C&E: inference about cause and effect

- 有时experiment不能完成因为不现实，因此我们只能通过observational study to make conclusions

因此，要想conclude，满足

① association is strong

②. association is consistent

③. larger values of explanatory variable are associated with stronger response

④ the alleged cause precedes the effect in time

⑤. the alleged cause is plausible

17. - point estimator

a statistic that provides an estimate of a population parameter. the value is point estimate

- confidence interval

we are C% confident that the interval from __ to __ capture, the ______

- CI for p: $\hat{p} \pm Z_{critical}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

conditions

① data comes from a well-design random sample or randomized experiment

② if without replacement, $n \leq \frac{1}{10}N$

③ both $n\hat{p}$, $n\hat{q} \geq 10$

sometimes 我们要确定n大小

$Z^*_{critical}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq ME_{MAX} < ME_{want}$

where ME_{MAX}: maximum margin of error occurs when $p = 0.5$

- standard error

when the standard deviation of a statistic is measured from the data.

randomized experiment

② if without replacement. $n \le \frac{1}{10}N$

③ population has a normal distribution or $n \geq 30$; if distribution unknown and $n < 30$, use a graph of sample data to assess the normality of population. don't use t procedure if graph shows skewness or outliers

sometimes, need to find n first

$Z_{critical}\frac{\sigma}{\sqrt{n}} \le ME_{want}$

其中σ是先前实验的结果

19. • differece between p

① for p_1-p_2:

$\mu_{\hat p_1-\hat p_2} = p_1 - p_2$

$\sigma_{\hat p_1-\hat p_2} = \sqrt{\sigma_{\hat p_1}^2+\sigma_{\hat p_2}^2} = \sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}$

$\hat p_1-\hat p_2$ is approximately normal if

$n_1p_1,\ n_1(1-p_1),\ n_2p_2,\ n_2(1-p_2) \geq 10$

② $CI = (\hat p_1-\hat p_2) \pm z_{critical}\sqrt{\frac{\hat p_1(1-\hat p_1)}{n_1}+\frac{\hat p_2(1-\hat p_2)}{n_2}}$

$$Z_{critical} = \frac{(\hat p_1-\hat p_2)-0}{\sqrt{\frac{\hat p_c(1-\hat p_c)}{n_1}+\frac{\hat p_c(1-\hat p_c)}{n_2}}}$$

where $\hat p_c = \frac{X_1+X_2}{n_1+n_2}$

(pooled sample proportion)

- difference for μ

① for $\mu_1-\mu_2$:

$\mu_{\bar x_1-\bar x_2} = \mu_1-\mu_2$.

$\sigma_{\bar x_1-\bar x_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

$\bar x_1-\bar x_2$ is approximately normal if

$n_1, n_2 \geq 30$

② $t_{critical} = \frac{(\bar x_1-\bar x_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

use t distribution with degrees of freedom equal to min (n_1-1, n_2-1)

$CI: (\bar x_1-\bar x_2) \pm t_{critical}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$

[paired只在一组人均经历了两种treatments时使用]

18. 假设检验

- P-value：如果H_0成立，那么发生该事件或更极端事件概率
- statistically significant at level α：P-value 小于 α
- 在α significant level的假设检验中
  $\alpha = P(\text{Type I error} \mid H_0 \text{ is true})$
- power of a test against a specific alternative is the probability that the test will reject H_0 at a chosen significant level α when the specified alternative value of the parameter is true
- P(Type I)与P(Type II)，一个下降则另一个上升

increase power by:

① increase sample size

② increase α

③ increase the difference that is important to detect between null and alternative parameter values

- 对P检验用z，对μ检验用t
  (自由度n-1，找不到就向下取整)
- 影响size增加因素

① smaller significant level: need stronger evidence to reject H_0

② higher power: higher power gives a better chance of detecting a difference when it exits

③ smaller effect size

注：significant ≠ important

20. • chi-square distribution

的自由度为n [illegible] n-1

mean = n

mode = n-2

condition:

① well-designed random sample or randomized experiment.

② $n \le \frac{1}{10}N$

③ all expected counts ≥ 5

reject H_0，最大的component ($\frac{(O_i-E}{E_i}$

21 regression

- choose an SRS of n observations (a,b) from a population of size N with least-squares regression line (population)

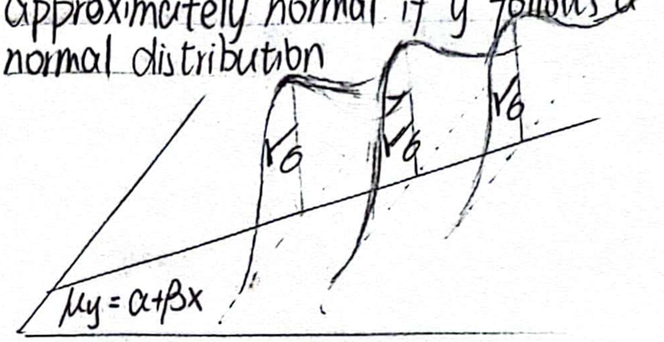
predict $y = \alpha + \beta x$

let b be the slope of the sample regression line. then

$\mu_b = \beta$ $\sigma_b = \frac{\sigma}{\sigma_x \sqrt{n}}$ as long as $n \le \frac{1}{10} N$

where σ_x is the SD of N observations.

the sample distribution of b is approximately normal if y follows a normal distribution



- condition.

① linear: use scatter plot, centred at residual=0

② independent

if without replacement, $n \le \frac{1}{10} N$

③ normal stemplot/histogram/normal probability plot

④ equal SD: no matter value of x.

look at the scatter of the residuals above and below the 'residual=0' line: vertical spread should be same no matter the value of x

⑤ random

- CI of $\beta = b \pm t_{critical} \frac{s}{s_x \sqrt{n-1}}$

where $s = \sqrt{\frac{\Sigma (y_i - \hat{y}_i)^2}{n-2}}$

$s_x = \sqrt{\frac{\Sigma (x - \bar{x})^2}{n-1}}$

t的自由度为n-2

- $t = \frac{b - \beta}{\frac{s}{s_x \sqrt{n-1}}}$

t的自由度为n-2.

判断有无linear关系, $H_0 = \beta = 0$

22 transforming to achieve linearity

① power model $y = ax^p \Rightarrow (x^p, y) / (x, \sqrt[p]{y})$

② exponential model. $y = ab^x \Rightarrow (x, \ln y)$

AP 微积分

1. δ与ε语言

2. 夹逼. 通分有理化 (if $g(x)=f(x)$ for $x\neq c$, then $\lim_{x\to c} f(x)=\lim_{x\to c} g(x)$)

3. (derivative) differentiability → continuity，不能反推

4. **rolle's theorem**: let $f(x)$ be continuous on the closed interval $[a,b]$, and differentiable on the open interval (a,b). if $f(a)=f(b)$. then there is at least one number c in (a,b) such that $f'(c)=0$

mean value theorem: if $f(x)$ is continuous on the closed interval $[a,b]$ and differentiable on the open interval (a,b). then there exists a number c in (a,b) such that $f'(c)=\frac{f(b)-f(a)}{b-a}$

Proof: line passing through $(a,f(a))$, $(b,f(b))$

$y=\left(\frac{f(b)-f(a)}{b-a}\right)(x-a)+f(a)$ [注意 $\frac{f(b)-f(a)}{b-a}$ 与 $f(a)$ 为 constant]

let $g(x)$ be the difference between $f(x)$ and y

$g(a)=g(b)=0$, since $f(x)$ is continuous, so does $g(x)$

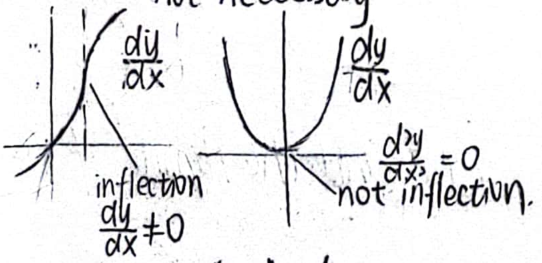
exist c on (a,b) satisfying $g'(c)=0$

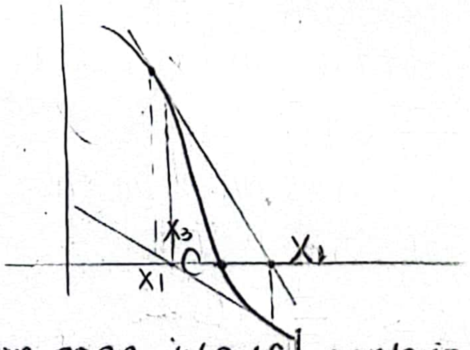
$g(x)=f(x)-y=f(x)-\left(\frac{f(b)-f(a)}{b-a}\right)(x-a)-f(a)$

$g'(x)=f'(x)-\frac{f(b)-f(a)}{b-a}$ $\quad g'(c)=0 \quad f'(c)=\frac{f(b)-f(a)}{b-a}$

5. inflection: the point when concave upwards (downwards) turns into downwards (upwards)

$$\frac{d^2y}{dx^2}=0 \underset{\text{not neccessary}}{\overset{\text{not neccessary}}{\rightleftarrows}} \text{inflection} \underset{\text{must be}}{\overset{\text{not neccessary}}{\rightleftarrows}} \text{turning point (max/min)}$$





6. **Newton's method of approximation**

let $f(c)=0$, where $f(x)$ is differentiable on an open interval containing c

we want to find an approximation of c [the bound[illegible] value in

- make an initial estimate x_1
- $x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$, again and again.

because: tangent passing though x_1: $y-f(x_1)=f'(x_1)(x-x_1)$

$y=f'(x_1)(x-x_1)+f(x_1)$

when $y=0$ $\quad x=x_1-\frac{f(x)}{f'(x_1)}$

into to successfully use this method $\left|\frac{f(x)f''(x)}{[f'(x)]^2}\right|<1$.

7. • when $\Delta x \to 0$, $\Delta y \to dy$:

$\Delta y = f(x+\Delta x) - f(x)$, $dy = f'(x)dx$,

• find the area (integral) of graph in $[a,b]$ (不是定义)

Δx (步长) $= \frac{a-b}{n}$, $A = \sum_{i=1}^{n} f(a+n\Delta x)\cdot\Delta x$

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(a+n\Delta x)\Delta x$$

[可以设每一项取 upper/lower limit 来确定 boundaries. 当 $n\to\infty$ 时每一项的 upper limit = lower limit]

let f be continuous and nonnegative on the interval $[a,b]$; the limits as $n\to\infty$ of both the lower and upper sums exist and are equal to each other.

$$\lim_{n\to\infty} s(n) = \lim_{n\to\infty} \sum_{i=1}^{n} f(m_i)\Delta x = \lim_{n\to\infty} \sum_{i=1}^{n} f(M_i)\Delta x = \lim_{n\to\infty} S(n)$$

where $\Delta x = \frac{b-a}{n}$ and $f(m_i)$ and $f(M_i)$ are the minimum and maximum value of f on the subinterval.]

• let f be defined on the closed interval $[a,b]$, and let Δ be a partition of $[a,b]$ given by

$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$

where Δx_i is the width of the ith subinterval. if c_i is any point in the ith subinterval $[x_{i-1}, x_i]$, then the sum (中值定理)

$\sum_{i=1}^{n} f(c_i)\Delta x_i$, $x_{i-1} \le c_i \le x_i$

is called the Riemann sum of f for the partition Δ

the width of the largest subinterval of a partition Δ is the norm of the partition and is devoted by $\|\Delta\|$

$\frac{b-a}{\|\Delta\|} \le n$.

• if f is defined on the closed interval $[a,b]$ and the limit of Riemann sums over partition Δ

$$\lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i$$

exists, then f is said to be integrable on $[a,b]$ and the limit is denoted by

$$\lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(c_i)\Delta x_i = \int_a^b f(x)dx$$

(可用 ε 与 δ 语言表示)

• If a function f is continuous on the closed interval $[a,b]$ and F is an antiderivative of f on the interval $[a,b]$, then

$$\int_a^b f(x)dx = F(b) - F(a)$$

proof. let Δ be any partition of $[a,b]$

$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$

$\therefore F(b) - F(a) = F(x_n) - F(x_{n-1}) + F(x_{n-1}) - F(x$

$$= \sum_{i=1}^{n} [F(X_i) - F(X_{i-1})]$$

by mean value theorem, there exists a number c_i in the ith subinterval.

$$F'(c_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}$$

$\because x_i - x_{i-1} = \Delta x$ and $F'(c_i) = f(c_i)$

$$F(b) - F(a) = \sum_{i=1}^{n} f(c_i)\Delta x_i$$

- $\frac{d}{dx}[\int_a^x f(t)dt] = f(x)$

proof. defining $F(x) = \int_a^x f(t)dt$

$$F'(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{1}{\Delta x} [\int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt]$$

$$= \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_x^{x+\Delta x} f(t)dt$$

by mean value theorem, we can find a value of c: $\int_x^{x+\Delta x} f(t)dt = f(c)\Delta x$

$$F'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x} f(c)\Delta x = \lim_{\Delta x \to 0} f(c)$$

$\because x \leq c \leq x+\Delta x$

$\therefore F'(x) = \frac{d}{dx}[\int_a^x f(t)dt] = f(x)$

8. find the approximation of an integral: Simpson's Rule

if $p(x) = Ax^2 + Bx + C$

$$\int_a^b p(x)dx = \int_a^b Ax^2 + Bx + C\, dx = [\frac{Ax^3}{3} + \frac{Bx^2}{2} + Cx]_a^b$$

$$= \frac{A}{3}(b^3 - a^3) + \frac{B}{2}(b^2 - a^2) + C(b-a)$$

$$= \frac{b-a}{6}[2A(a^2+ab+b^2) + 3B(b+a) + 6C]$$

$$= \frac{b-a}{6}[p(a) + 4p(\frac{a+b}{2}) + p(b)]$$

We partition interval $[a,b]$ into n subinterval, each of width $\Delta x = \frac{b-a}{n}$ where n is an even, and the subintervals are grouped in pairs such that

$$a = \underbrace{x_0 < x_1 < x_2}_{[x_0, x_2]} \underbrace{< x_3 < x_4}_{[x_2, x_4]} < \cdots < \underbrace{x_{n-2} < x_{n-1} < x_n}_{[x_{n-2}, x_n]} = b$$

on each double subinterval $[x_{i-2}, x_i]$, we can approximate f by a polynomial p of degree less than or equal to 2.

$$\int_{x_0}^{x_2} f(x)dx \approx \int_{x_0}^{x_2} p(x)dx = \frac{x_2 - x_0}{6}[p(x_0) + 4p(\frac{x_0+x_2}{2}) + p(x_2)]$$

$$= \frac{b-a}{3n}[f(x_0) + 4f(x_1) + f(x_2)]$$

$$\therefore \int_a^b f(x)dx \approx \frac{b-a}{3n}[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots 4f(x_{n-1}) + f(x_n)]$$

if f has a continuous second derivative on $[a,b]$, then the error E in approximating $\int_a^b f(x)dx$ by the Trapezoidal Rule is

$$|E| \leq \frac{(b-a)^3}{12n^2}[\max|f''(x)|], \quad a \leq x \leq b$$

if f has a continuous fourth derivative on $[a,b]$, then the error E in approximating $\int_a^b f(x)dx$ by Simpson's Rule is

$$|E| \leq \frac{(b-a)^5}{180n^4}[\max|f^{(4)}(x)|], \quad a \leq x \leq b$$

9. let f be a function that is differentiable on an interval I, if f has an inverse function g, then g is differentiable at any x for which $f'(g(x)) \neq 0$, moreover

$$g'(x) = \frac{1}{f'(g(x))}, \quad f'(g(x)) \neq 0. \quad \left(\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}\right)$$

10. $\frac{d}{dx}(a^x) = \ln a \cdot a^x$ $\quad$ $\frac{d}{dx}(\log_a x) = \frac{1}{x\ln a}$ $\quad$ $\lim_{x\to\infty}(1+\frac{1}{x})^x = e$

$\frac{d}{dx}\operatorname{arccot}x = -\frac{1}{1+x^2}$ $\quad$ $\frac{d}{dx}\operatorname{arcsec}x = \frac{1}{|x|\sqrt{x^2-1}}$ $\quad$ $\frac{d}{dx}\operatorname{arccsc}x = -\frac{1}{|x|\sqrt{x^2-1}}$

$\int \frac{1}{x\sqrt{x^2-a^2}}dx = \frac{1}{a}\operatorname{arcsec}\frac{|x|}{a} + C$

$\coth^{-1}x = \frac{1}{2}\ln\frac{x+1}{x-1}$ $\quad$ $\operatorname{sech}^{-1}x = \ln\frac{1+\sqrt{1-x^2}}{x}$ $\quad$ $\operatorname{csch}^{-1}x = \ln\left(\frac{1}{x} + \frac{\sqrt{1+x^2}}{|x|}\right)$

$\int \frac{1}{\sqrt{u^2\pm a^2}}du = \ln(u+\sqrt{u^2\pm a^2})$ $\quad$ $\int \frac{1}{a^2-u^2}du = \frac{1}{2a}\ln\left|\frac{a+u}{a-u}\right|$ $\quad$ $\int \frac{1}{u\sqrt{a^2\pm u^2}}du = -\frac{1}{a}\ln\frac{a+\sqrt{a^2\pm u^2}}{|u|}$

11. euler's method of approximating the particular solution of DE

for $y' = F(x,y)$ and a small value of h

start with (x_0, y_0): $\quad x_1 = x_0 + h, \quad y_1 = y_0 + hF(x_0, y_0)$

$\quad x_2 = x_1 + h, \quad y_2 = y_1 + hF(x_1, y_1)$

12. homogeneous differential equation of solving DE

if $f(tx, ty) = t^n f(x,y)$, then $f(x,y)$ is homogeneous of degree n.

a homogeneous differential equation is an equation is the form of

$M(x,y)dx + N(x,y)dy = 0$

where $M(x,y)$ and $N(x,y)$ are homogeneous functions of same degree

in this way, the equation can be transformed into a differential equation whose variables are seperable by substitution

$y = vx$, where v is differentiable function of x

- 在解DE时，我们一般会写出一组解，我们称这样的一组解 family of curv
  两个families of curves 如果intersection时一点都normal，我们称
  mutually orthogonal
  each curve in one of the families is called an orthogonal trajectory of the other family

A rational function is the quotient of two polynomial functions.

PS注：这一块N(t)的推导来源于，if $\vec{r}(t)\cdot\vec{r}(t)=c$, then $\vec{r}(t)\cdot\vec{r}'(t)=0$.

if a convergent alternating series

求解orthogonal trajectories:

if $\frac{dy}{dx}=F(x,y)$，我们只要求 $\frac{dy}{dx}=-\frac{1}{F(x,y)}$ 的solution即可

- **Bernoulli Equation** $\frac{dy}{dx}+P(x)y=Q(x)y^n$

we only focus on the situation when $n\neq 0$ or 1

$$y^{-n}\frac{dy}{dx}+P(x)y^{1-n}=Q(x)$$

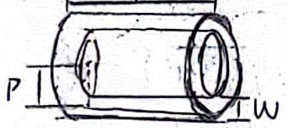
$$(1-n)y^{-n}\frac{dy}{dx}+(1-n)P(x)y^{1-n}=(1-n)Q(x)$$

$$\frac{d}{dx}[y^{1-n}]+(1-n)P(x)y^{1-n}=(1-n)Q(x)$$

let $z=y^{1-n}$: $\frac{dz}{dx}+(1-n)P(x)z=(1-n)Q(x)$

13. 用积分求体积/做功/CoG/净力

- $V=\int_a^b A\,dx$ [这里A的表达式可能与x有关]
- **shell method** 对于一个有hole的cylinder



p: average radius
h: height
w: thinkness

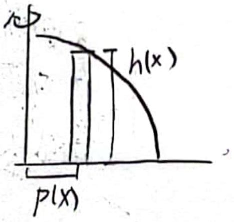
$$\Delta V=V_{cylinder}-V_{hole}$$
$$=\pi(p+\frac{w}{2})^2h-\pi(p-\frac{w}{2})^2h$$
$$=2\pi wph$$

绕x-axis转

$$V=2\pi\int_a^b p(y)h(y)dy$$

绕y-axis转

$$V=2\pi\int_a^b p(x)h(x)dx$$



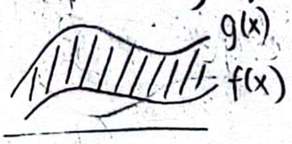
[注意此方法重点为找cylinder-hole，所以要注意p与h用x还是y表达]

- $W=\int F\,dx$

[注意，可以是$\Delta F\cdot x$与$F\cdot\Delta x$，要搞明白F与x]

气体：$W=\int\frac{k}{V}dV$, where $P=\frac{k}{V}$, P is pressure

-

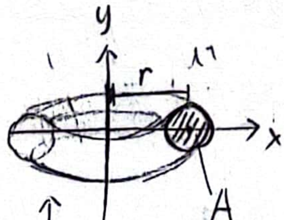


阴影的CoG

$$\bar{y}=\frac{\rho\int_a^b\overbrace{[\frac{f(x)+g(x)}{2}]}^{中心点}\overbrace{[f(x)-g(x)]}^{高}dx}{\int_a^b f(x)-g(x)dx}$$

$$\bar{x}=\frac{\rho\int_a^b x[f(x)-g(x)]dx}{\int_a^b f(x)-g(x)dx}$$

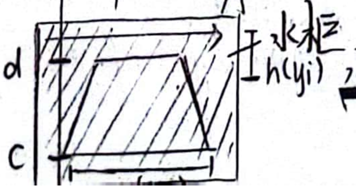
-



截面绕y轴旋转

$V=2\pi rA$, where A: 截面面积，
r: 截面中心至旋转轴距离

-



水柜内竖放物体受水的作用力

$F=w\int_c^d h(y)L(y)dy$, where w is weight-density)

$h(y)$ is the depth of fluid.
$L(y)$ is the horizontal length

14. 三角函数积分

- $\int \sin^{2k+1} x \cos^n x\,dx = \int \sin^{2k} x \cos^n x (\sin x\,dx)$
  $\int \sin^n x \cos^{2k+1} x\,dx = \int \sin^n x \cos^{2k} x (\cos x\,dx)$ $\Big\}$ $\sin^2 x + \cos^2 x = 1$
- if n is odd, $\int_0^{\frac{\pi}{2}} \cos^n x\,dx = \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \cdot (\frac{n-1}{n})$
  if n is even $(n \geqslant 2)$ $\int_0^{\frac{\pi}{2}} \cos^n x\,dx = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdots (\frac{n-1}{n}) \cdot \frac{\pi}{2}$
  similarly if turn $\cos^n x$ into $\sin^n x$
- $\int \sec^{2k} x \tan^n x\,dx = \int (\sec^2 x)^{k-1} \tan^n x (\sec^2 x\,dx)$
  $\int \sec^n x \tan^{2k+1} x\,dx = \int \sec^{n-1} x (\tan^2 x)^k (\sec x \tan x\,dx)$ $\Big\}$ $1+\tan^2 x = see^2 x$
  $\int \tan^n x\,dx = \int \tan^{n-2} x ((\sec^2 x - 1)dx)$
  $\int \sec^{2k+1} x\,dx$: use integration by parts
- 三角转化：
  for $\sqrt{a^2-u^2}$: $u = a\sin\theta$ $\int \sqrt{a^2-u^2}\,du = \frac{1}{2}(a^2 \sin^{-1}\frac{u}{a} + u\sqrt{a^2-u^2})$
  for $\sqrt{a^2+u^2}$ $u = a\tan\theta$ $\int \sqrt{u^2+a^2}\,du = \frac{1}{2}(u\sqrt{u^2+a^2} + a^2 \ln|u+\sqrt{u^2+a^2}|$
  for $\sqrt{u^2-a^2}$ $u = a\,see\theta$ $\int \sqrt{u^2-a^2}\,du = \frac{1}{2}(u\sqrt{u^2-a^2} - a^2 \ln|u+\sqrt{u^2-a^2}|$

15. 洛比达

如果variable在指数上, turn $y = f(x)$ into $\ln y = \ln f(x)$
对 $\ln y$ 洛后求 y

16.

- integral test 求series是否converge
  1. 保证 $f(x)$ is positive and continuous and decreasing from start
  2. 用inscribe证明 $(\int_1^{\infty} f(x)dx)$
- direct comparison test 求series是否converge/diverge
  if $0 < a_n < b_n$ for all n.
  if $\sum_{n=1}^{\infty} b_n$ converges, so does $\sum_{n=1}^{\infty} a_n$; if $\sum_{n=1}^{\infty} a_n$ diverges, so does $\sum_{n=1}^{\infty} b_n$
- limit comparison test 求series是否converge/diverge
  if $a_n > 0$, $b_n > 0$ for all n.
  $\lim_{n\to\infty} (\frac{a_n}{b_n}) = L$ where L is finite and positive
  then $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ both converge or diverge

  对于 $\frac{\sum_{n=0}^{N} a_n x^n}{\sum_{m=0}^{M} a_m x^m}$; limit 取 $\frac{x^N}{x^M}$ ，如果有 c^n 的话原位置copy
- series with alternative signs (alternative-sign test 求series converge)
  let $a_n > 0$, the $\sum_{n=1}^{\infty} (-1)^n a_n$ and $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$ converge if
  $\lim_{n\to\infty} a_n = 0$ and $a_{n+1} \leq a_n$ for all n

if a convergent alternating series satisfies the condition $a_{n+1} \le a_n$ then the absolute value of the ramainder R_N involved in approximating the sum S by S_N is less than (or equal to) the first neglect term,

$|S-S_N| = |R_N| \le a_{N+1}$

- if $\Sigma|a_n|$ converges, so does Σa_n
  Σa_n is absolutely convergent if $\Sigma|a_n|$ converges
  Σa_n is conditionally convergent if Σa_n converges but $\Sigma|a_n|$ diverges
- n-th term test 于series diverge if $\lim_{n\to\infty} a_n \ne 0$, $\sum_{n=1}^{\infty} a_n$ diverge
- p-series 于 series diverge or converge
  for $\sum_{n=1}^{\infty} \frac{1}{n^p}$: when $p>1$, converge, when $0 \le p \le 1$, diverge
- ratio test 于 series diverge or converge
  Σa_n converges absolutely if $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$
  Σa_n diverges if $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$ or $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty$
  ratio test is inconclusive if $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1$
- root test 于 series diverge or converge
  Σa_n converges absolutely if $\lim_{n\to\infty} \sqrt[n]{|a_n|} < 1$
  Σa_n diverges if $\lim_{n\to\infty} \sqrt[n]{|a_n|} > 1$ or $\lim_{n\to\infty} \sqrt[n]{|a_n|} = \infty$
  root test is inconclusive if $\lim_{n\to\infty} \sqrt[n]{|a_n|} = 1$
- strategy judging series divergent or convergent
  1. nth approach 0?
  2. special types? $\longrightarrow$ geometric, p-series, telescoping, alternating
  3. can integral test, root test, ratio test applied?
  4. can it be compared to one of special types?

17.
- taylor polynomial $f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c$
- malclaurin polynomial $f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n$
- for taylor polynomial $P_n(x)$, its error $R_n(x)$ is
  $|R_n(x)| = |f(x) - P_n(x)|$
  taylor theorem: if f is differentiable through order $n+1$ at interval I containing c, for each x in I, there exists z between x and c,
  $R_n(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x-c)^{n+1}$
  when n [illegible]

18. if x is a variable, then an infinite series of the form $\sum_{n=0}^{\infty} a_n x^n$ is called a power series. More generally, an infinite series of the form $\sum_{n=0}^{\infty} a_n (x-c)^n$ is called a power series centered at C

for a power series centered at C, there exists a real number $R>0$ such that the series converges absolutely for $|x-c|<R$, and diverge for $|x-c|>R$. R is called the radius of convergence and the set of all values of x for which series converges is interval of convergenc

- 求R：用test（ratio test居多）求出一个结果，然后看如果满足test的converge要求x要怎样

求interval of convergence：两端endpoint的converge情况要独立代入验证

- 将power series求导/积分后，原有的R不变，但endpoint的情况可能变。有时将复杂形式求积分/导数推interval
- geometric power series 改写

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}, \quad |r|<1$$ 硬凑形式即可

- operation with power series may change the interval of convergence（并集）

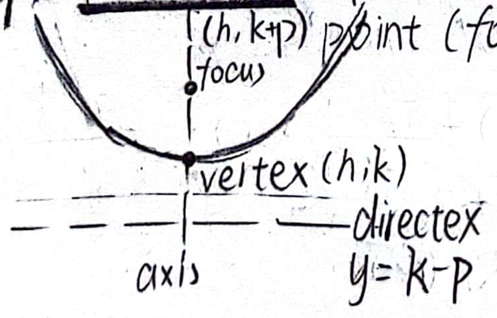
所以说对于一些复杂的geometric power series的分数形式，先partial，再分别求完interval后看并集

- talyor展开时，对于一些难求导的部分f(x)，先设g(x) = 简单部分(x

再换成f(g(x))即可

也可将其换为两个好算的展开的积/商

19. - parabola: equidistance from a fixed line (directex) and a fixed point (focus)



given that vertex is (h,k)

if directex is $y=k-p$, focus is $(h,k+p)$

curve: $(x-h)^2=4p(y-k)$

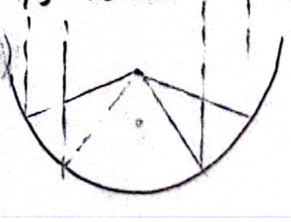
if directex is $x=h-p$, focus is $(h+p,k)$

curve: $(y-k)^2=4p(x-h)$

line segment passing through the focus of a parabola and has endpoints on the parabola is called focal chord.

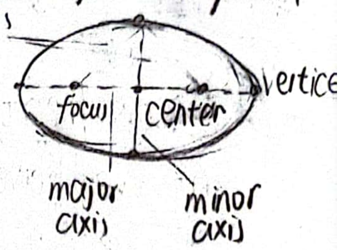
the special focal chord perpendicular to axis is latus rectum.

特殊性质：从focus到parabola上任意一点(x,y)，该点normal平分focus与该点和axis方向组成的角



L：王：这一块N(t)的推导来源于, if $\vec{r}(t)\cdot\vec{r}(t)=c$, then $\vec{r}(t)\cdot\vec{r}'(t)=0$,
即$|\vec{r}(t)|$不变 $\vec{r}$与$\vec{r}'$垂直

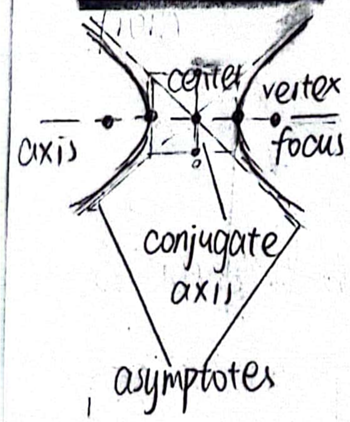
- **ellipse** sum of distances from one point to each of two fixed points (focus/foci) is the same
  if center is (h,k), the lengths of major and minor axes are $2a, 2b$, where $a>b$



  horizontal major axis: $\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1$
  vertical major axis: $\frac{(x-h)^2}{b^2}+\frac{(y-k)^2}{a^2}=1$
  the foci lies on major axis, c unit from center, where $c^2=a^2-b^2$

ellipse有与parabola一样的特殊性质
eccentricity: $e=\frac{c}{a}$ 面积用极坐标, 周长带公式

- **hyperbolas**: the abosulte value of difference between distance from one point to each of the fixed point (focus/foci) is the same



  if center is (h,k)
  horizontal axis: $\frac{(x-h)^2}{a^2}-\frac{(y-k)^2}{b^2}=1$
  vertical axis: $\frac{(y-k)^2}{b^2}-\frac{(x-h)^2}{a^2}=1$
  the vertices are a units from center and foci are c units from center, where $c^2=a^2+b^2$
  conjugate is the line segment of length 2b joining:
  (horizontal axis) $(h, k+b)$ $(h, k-b)$
  (vertical axis) $(h+b, k)$ $(h-b, k)$
  hyperbolas have two asymptotes
  (horizontal axis) $y=k\pm\frac{b}{a}(x-h)$
  (vertical axis) $y=k\pm\frac{a}{b}(x-h)$

eccentricity: $e=\frac{c}{a}$

20. **parametric equation** is smooth as long as $\frac{dx}{dt}$ and $\frac{dy}{dt}$ are continuous on the demand and not simultaneously 0, except endpoints.
curve is piecewise smooth if it is smooth on each subinterval of some partition.

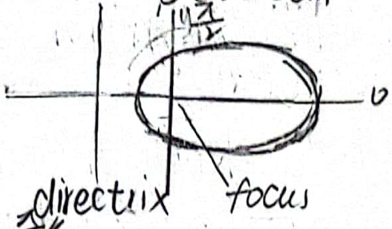
21. **polar equation**: $f(\theta)$ is a function in polar system
$\frac{dx}{d\theta}=0$, $\frac{dy}{d\theta}\neq0$: horizontal tangent
$\frac{dy}{d\theta}=0$ $\frac{dx}{d\theta}\neq0$: vertical tangent

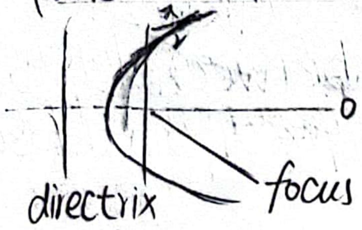
- 求两(polar system) curves 相交时，考虑 $(r(\theta), \theta)$ $(y(-\theta), -\theta)$ [illegible] 点

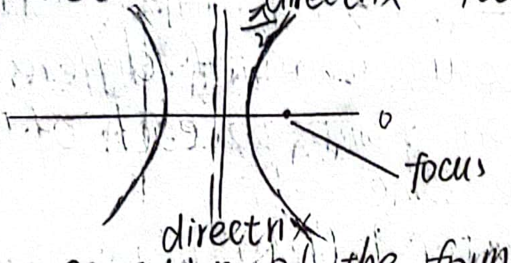
22. classification of conics by eccentricity:

if we have a fixed point (focus) and a fixed line (directrix). let eccentricity (e) be the ratio of the distance between focus and the point to the distance between directrix to the point

conic $\begin{cases} \text{when } 0<e<1 & \text{ellipse} \\ \text{when } e=1 & \text{parabola} \\ \text{when } e>1 & \text{hyperbola} \end{cases}$



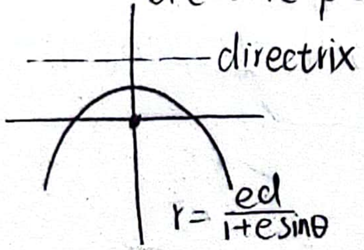


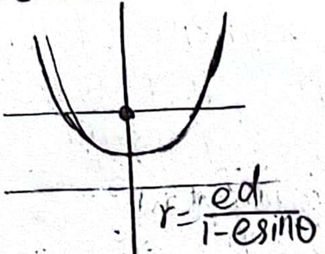


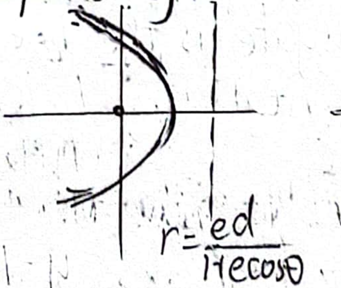
the graph of a polar equation of the form

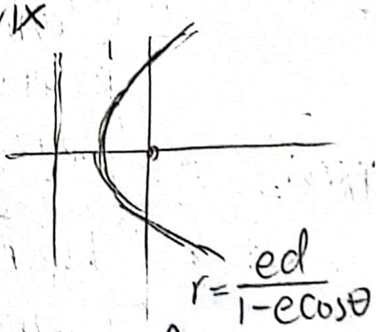
$$r = \frac{ed}{1 \pm e\cos\theta} \quad \text{or} \quad r = \frac{ed}{1 \pm e\sin\theta}$$

is a conic, where $e>0$ and $|d|$ is the distance between the focus at the pole and its corresponding directrix









ellipse:

$$b^2 = a^2 - c^2 = a^2 - (ea)^2 = a^2(1-e^2)$$

since $e = \frac{c}{a}$

hyperbola

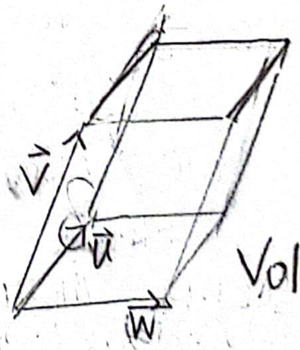
$$b^2 = c^2 - a^2 = a^2(e^2-1)$$

23. vector

$\vec{v} \cdot (\vec{u} + \vec{w}) = \vec{v} \cdot \vec{u} + \vec{u} \cdot \vec{w}$

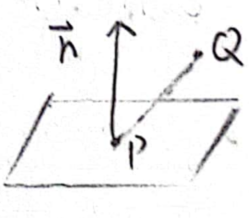
$\vec{v} \times (\vec{u} + \vec{w}) = \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{w}$

$\vec{v} \cdot (\vec{u} \times \vec{w}) = (\vec{u} \times \vec{w}) \cdot \vec{v}$



Volume $= \vec{v} \cdot (\vec{u} \times \vec{w})$

distance from a point to the plane: $\frac{|\vec{PQ} \cdot \vec{n}|}{|\vec{n}|}$



24. 三维坐标系中图形表达

sphere: $(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 = r^2$

cylinders: 只于两个 variables 有关，在另一方向上可无限延伸.

quadric surface:

[上述 [illegible] 属于, if $\vec r(t)\cdot\vec r(t)=c$, then $\vec r(t)\cdot\vec r'(t)=0$,
即$|\vec r(t)|$不变, $\vec v$与$\vec a$垂直]

- **$\vec a$与$\vec T(t)$ $\vec N(t)$关系**

$$\vec T(t)=\frac{\vec r'(t)}{|\vec r'(t)|}=\frac{\vec v}{|\vec v|} \qquad \vec N(t)=\frac{\vec T'(t)}{|\vec T'(t)|}$$

$$\vec a=\vec v'=[\vec T(t)|\vec v|]'=(|\vec v|)'\vec T(t)+|\vec v|\vec T'(t)$$
$$=(|\vec v|)'\vec T(t)+\underline{|\vec v||\vec T'(t)|}\vec N(t)$$
$$=a_T\vec T(t)+a_N\vec N(t)$$

$$a_T=(|\vec v|)'=\vec a\cdot\vec T(t)=\frac{\vec v\cdot\vec a}{|\vec v|}$$
$$a_N=|\vec v||\vec T'(t)|=\vec a\cdot\vec N(t)=\frac{|\vec v\times\vec a|}{|\vec v|}=\sqrt{|\vec a|^2-a_T^2}$$

note that $a_N\geqslant 0$ and normal component of acceleration is called centripetal component of acceleration.

- **expressing arc length using vector** $S=\int_a^b|\vec r'(t)|dt$

arc length parameter $s(t)=\int_a^t|\vec r'(u)|du,\ a\leq t\leq b$

if s is arc length parameter and $s=f(t)$, convert $\vec r_1(t)$ into $\vec r_2(s)$ then $|\vec r_2'(s)|=1$. if t is any parameter for the vetor-valued funtion $\vec r$ such that $|\vec r'(t)|=1$, then t must be arc length parameter

curvature $K=|\frac{d\vec T}{ds}|=|\vec T'(s)|$ $\qquad K=\frac{|\vec T'(t)|}{|\vec r'(t)|}=\frac{|\vec r'(t)\times\vec r''(t)|}{|\vec r'(t)|^3}$ $=\frac{\vec a(t)\cdot\vec N(t)}{|\vec v(t)|}$

[proof: $\frac{\vec T'(t)}{ds/dt}\approx\frac{[\vec T(t+\Delta t)-\vec T(t)]/\Delta t}{[s(t+\Delta t)-s(t)]/\Delta t}=\frac{\Delta T}{\Delta s}$]

if C is the graph of a twice-differentiable function given by, then $K=\frac{|x'y''-y'x''|}{[(x')^2+(y')^2]^{\frac{3}{2}}}$ [proof by calculation]

let C be a curve with curvature K at point P, the circle passing through P with radius $r=\frac{1}{K}$ is called the circle of curvature. the radius of curvature at P is r and the center is called the center of curvature

$$\vec a=(|\vec v|)'\vec T(t)+|\vec v||\vec T'(t)|\vec N(t)=\frac{d^2s}{dt^2}\vec T(t)+\frac{ds}{dt}(|\vec v|K)\vec N(t)$$
$$=\frac{d^2s}{dt^2}\vec T(t)+K(\frac{ds}{dt})^2\vec N(t)$$

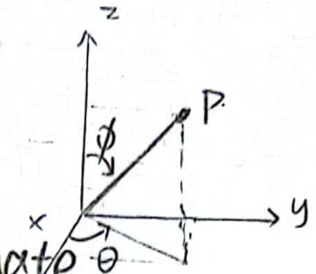
- **functions of several variables**

if a function that can be written as a sum of functions of the form cx^my^n (c is a real number and m and n are non-negative) is called a polynomial function of two variable revolution: 绕x轴 $y^2+z^2=[r(x)]^2$

绕y轴 $x^2+z^2=[r(y)]^2$

绕z轴 $x^2+y^2=[r(z)]^2$

反推时找相同的二次项系数



25. cylindrical coordinates (r,θ,z)

spherical coordinates (ρ,θ,ϕ)

ρ: radius θ: similar to cylindrical coordinate

ϕ: angle between the positive z-axis to $\overrightarrow{OP}$, where P is the poin

$0\leq\phi\leq\pi$

转化至rectangular coordinate

$x=\rho\sin\phi\cos\theta$ $y=\rho\sin\phi\sin\theta$ $z=\rho\cos\phi$

$\rho=\sqrt{x^2+y^2+z^2}$ $\theta=\tan^{-1}\frac{y}{x}$ $\phi=\cos^{-1}\frac{z}{\rho}$

转化至cylindrical coordinates

$r=\rho\sin\phi$ $\theta=\theta$ $z=\rho\cos\phi$

$\rho=\sqrt{r^2+z^2}$ $\phi=\cos^{-1}\frac{z}{\rho}$

26. ~~vector-valued function~~ ($r(t)$ and $u(t)$ are vector-valued function)

$\vec{r}(t)=f(t)\vec{i}+g(t)\vec{j}+h(t)\vec{k}$

$\vec{r}'(t)=f'(t)\vec{i}+g'(t)\vec{j}+h'(t)\vec{k}$

$[\vec{r}(t)\cdot\vec{u}(t)]'=\vec{r}'(t)\cdot\vec{u}(t)+\vec{r}(t)\cdot\vec{u}'(t)$

$[\vec{r}(t)\times\vec{u}(t)]'=\vec{r}'(t)\times\vec{u}(t)+\vec{r}(t)\times\vec{u}'(t)$

$[\vec{r}(w(t))]'=\vec{r}'(w(t))w'(t)$

If $\vec{r}(t)\cdot\vec{r}(t)=c$, then $\vec{r}(t)\cdot\vec{r}'(t)=0$ [$|\vec{r}(t)|$不变, 则$\vec{a}$与$\vec{v}$垂直]

$\int\vec{r}(t)dt=[\int f(t)dt]\vec{i}+[\int g(t)dt]\vec{j}+[\int h(t)dt]\vec{k}$

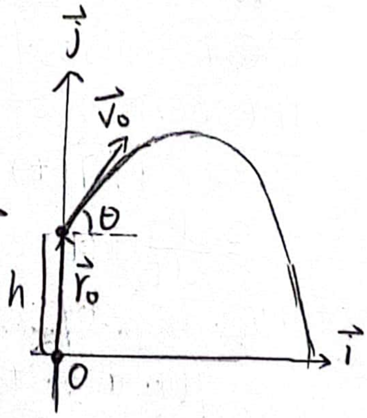
[注: 不定积分要加 $C=C_1\vec{i}+C_2\vec{j}+C_3\vec{k}$]

- 用vector-valued function解运动学

projectile motion中: $\vec{a}=-g\vec{j}$

position vector通式: $\vec{r}(t)=-\frac{1}{2}gt^2\vec{j}+t\vec{v_0}+\vec{r_0}$

$=(v_0\cos\theta)t\vec{i}+[h+(v_0\sin\theta)t-\frac{1}{2}gt^2]\vec{j}$



- unit tangent vector: $\vec{T}(t)=\frac{\vec{r}'(t)}{|\vec{r}'(t)|}$

unit normal vector: $\vec{N}(t)=\frac{\vec{T}'(t)}{|\vec{T}'(t)|}$

if $\vec{T}(t)=x(t)\vec{i}+y(t)\vec{j}$, then $\vec{N}(t)$ is either $y(t)\vec{i}-x(t)\vec{j}$ or $-y(t)\vec{i}+x(t)\vec{j}$

a rational function is the quotient of two polynomial functions

- **scalar field 标量场** can be characterized by level curves/contour lines 等高线

level curves 可以被扩展到三维 level surface: $f(x,y,z)=c$

- **interior point**: δ-neighborhood about (x_0,y_0) can be a disk centered at (x_0,y_0) with radius $\delta>0$. every δ-neighborhood about (x_0,y_0) lies entirely in R then (x_0,y_0) is interior point

boundary point: for points inside the disk, some are inside the R and some are not

open region: only contains all its interior points

closed region: contains all its interior and boundary points

- let f be a function of two variables defined, except possibly at (x_0,y_0) on an open disk centered at (x_0,y_0), and let L be a real number, then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)=L$

  if for each $\varepsilon>0$, there corresponds a $\delta>0$ such that $|f(x,y)-L|<\varepsilon$ whenever $0<\sqrt{(x-x_0)^2+(y-y_0)^2}<\delta$

  如果从任意方向上看结果都趋近一个值，则极限存在，反之则不然。

  continuous at point (x_0,y_0) if $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)=f(x_0,y_0)$

  polynomial and rational functions are all continuous at their domain.

① $f(x,y)$ $f_x(x,y)$ [illegible]

[illegible] $f_{yx}(x,y)$ mixed partial [illegible]

If f is a function of x and y such that f_x and f_y are continuous on the open disk R, then for every (x,y) in R $f_{xy}(x,y)$ [illegible] $f_{yx}(x,y)$

若 $z=f(x,y)$, $dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy$ [illegible]

[illegible] $=$ [illegible] $\Delta x+\frac{\partial z}{\partial y}\Delta y+\varepsilon_1\Delta x+\varepsilon_2\Delta y$ where [illegible] $\varepsilon_2\to0$ as $\Delta x\to0$ $\Delta y\to0$, then we consider z as differen[illegible]

If $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are continuous in an [illegible] then [illegible] is [illegible] differentiable on R

If [illegible] is differentiable at (x_0,y_0) [illegible]

- chain [illegible] variable [illegible] $\frac{\partial w}{\partial s}$

If $F(x,y)$ [illegible] then $\frac{dy}{dx}$ [illegible] $\frac{F_x(x,y)}{F_y(x,y)}$ where $F_y(x,y)$ [illegible]

If $F(x,y,z)$ [illegible] then $\frac{\partial z}{\partial x}=$ [illegible] $\frac{F_x(x,y,z)}{F_z(x,y,z)}$, $\frac{\partial z}{\partial y}=$ [illegible] $\frac{F_y(x,y,z)}{F_z(x,y,z)}$ where $F_z(x,y,z)$ [illegible]