Analysis of functions over fields: Quick reference

vishvAs vAsuki

April 16, 2012

Contents

1

IV Bounding techniques

See [1].

Part I

Introduction

1   Themes

Investigating C, functions of complex numbers.

2   Important Fields

2.1   General properties

For algebraic properties etc.., see algebra ref.
For Open balls, interior points, connectedness etc.. See topology ref.

2.2   Numbers

Natural numbers. Integers.

2.2.1   Supremum (GUB) vs infimum (GLB)

See properties of ordered sets in algebra ref.

2.2.2   The field of Rational numbers

Q; irrationality of $\sqrt{2}$; $\sqrt{2}$ as limit of rationals.
Q is an ordered field. Q is countable, as Z^2 is countable.

2.2.2.1   Absence of least upper bound

$\{x \in Q, x^2 < 2\}$ does not have supremum in Q.

2.3   The field of Real numbers

2.3.1   Supremum property

R is a field containing Q, with the supremum property. Every $S \subset R$ has supremum, infemum.

2.3.2   As Cuts from Q

Set elements of R as (Dedekind) cuts: $a \subset Q$ which split Q into 2 parts in any way such that: if $x \in a$, any y: $x > y \in a$, it has some upper bound in Q, it has no max element. Eg: $\{x \in Q, x^2 < 2\}$. Define $a < b$ by $a \subset b$, $a + b$ as $\{u + v | u \in a, v \in b\}$ etc; map $q \in Q$ to $(-\infty, q)$.

2.3.3 Archimedian property of R

If $0 < x < y$, $\exists n \in N : y < nx$. Else, take LUB l of $\{nx\}$, so $\exists n : l - x < nx$, then $l < (n+1)x$: $\Rightarrow\Leftarrow$.

2.3.4 Misc Properties

Decimal representation; its weakness: $.\bar{9} = 1$. There is a real between any 2 rationals. So: There is a rational between any 2 reals: also by archimedian property.
There is a real number between any two rationals: scale $\sqrt{2}$ as needed.

2.3.5 Extended real number system

$R \cup \{\infty, -\infty\}$: not a field.

2.3.6 Irrationals

Irrational numbers. Algebraic numbers: roots of polynomials with rational/ integer coefficients. Transcendental numbers: non algebraic; eg: e, π. R dense in irrationals: Rationals are countable, but R is not. R dense in transcendentals: algebraic numbers countable.

2.3.7 Topological properties k cell in R^k

Closed interval in R; Extend to hypercuboid in R^k. Also see topology ref.

2.3.7.1 Nonemptiness of telescoping $\cap$ of cells

If $\{I_n = [a_n, b_n]\}$ closed intervals in R with $I_n \supset I_{n+1}$ (as if diminishing), $\cap_j I_j \neq \phi$: take $E = \{a_n\}$; E bounded above by every b_n; so has sup E = x; this is in $\cap I_n$.

2.3.7.2 Compactness of k-cells in R^k

Take k-cell K definted by k closed intervals. K compact: For $\Rightarrow\Leftarrow$, take open cover $\{G_i\}$ sans finite subcover. Halve these intervals to break it into 2^k k-cells. If K is not compact: atleast one of these pieces, K_1 is not compact. Repeat the procedure with K_1 recursively ad infinitum. But, by theorem about diminishing intervals I_n, $\exists p \in \cap I_n$; as p interior pt of G_i, $\exists r, i : N_r(p) \supset K_i$: so K_i has finite cover: $\Rightarrow\Leftarrow$.
So, Bounded and closed intervals in R are compact.

2.3.7.3 Compactness in R^k

S closed and bounded $\equiv$ S Compact $\equiv$ C: Every ∞ subset E of S has a limit point of S: If S closed and bounded, it is in some k-cell, so compact. If S

compact, C is true. If C true: S is closed by defunction. If C true: S is bounded: else $\Rightarrow\Leftarrow$: $\{x_n : |x_n - x_0| > n\}$ wouldn't have lim pt in S.
So, compact set in R has supremum in it.

2.3.8 Uncountability of R

∞ sequences of $\{0,1\}$ uncountable by Cantor's diagonalization. So, $[0, 1]$ uncountable. So, looking at decimal representations, R is uncountable.
If $P \subset R^k, P \neq \phi$, P closed, every p in P is a limit point of P, then P uncountable: By a sort of diagonalization: As P has limit pts, P is ∞. For $\Rightarrow\Leftarrow$ let sequence $(x_i) = P$; take $V_1 = N_r(x_1) \cap P \neq \phi$; take $cl(V_1)$: is compact; make $V_2 = N_r(p) \cap V_1 \neq \phi$ with $x_1 \notin V_2$; do recursively; $\cap cl(V_i) \neq \phi$, but $\cap cl(V_i) \cap P = \phi :\Rightarrow\Leftarrow$.

2.3.8.1 Cantor set

Take $E_0 = [0, 1]$; remove middle third; do this recursively; $P = \cap E_i$. P is bounded and closed: so compact; also every pt is a lt pt: so uncountable; has no interval!
P has 0 measure. [**Find proof**]

2.3.8.2 Cardinality of R^k is same as R

Take $p = (a_1, ..a_k) \in R^k$, represent it in binary system; map it to x such that $x_{1:k}$ has first bits of p, $x_{k+1:2k}$ has second bits of p, etc..

2.4 The field of Complex numbers C

Set of ordered pairs over R, with +, * redifined to form a field: more than the complex plane R^2. An algebraic closure of R due to fundamental theorem of algebra.

2.4.1 Spherical form

$z = a + ib = r(\cos\theta + i\sin\theta)$.
Use Taylor series around 0: $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} \cdots = \frac{e^{xi}-e^{-xi}}{2i}$; $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \cdots = \frac{e^{xi}+e^{-xi}}{2}$. So, $(\cos\theta + i\sin\theta) = e^{i\theta}$.
So, $z = re^{i\theta}, y = r'e^{i\alpha}$ related by $y = \frac{r'}{r}e^{i(\alpha-\theta)}$.

2.4.2 Powers of z

(De Moivre). So, (Also by induction) $[r(cos\theta+isin\theta)]^n = r^n e^{ni\theta} = r^n(\cos n\theta + i\sin n\theta)$. So, $\bar{z}^n = \bar{z}^n$.

2.4.3 Other Properties

Magnitude $|z| = \sqrt{a^2 + b^2}$. $\sqrt{a + ib} = e + ih$. C has multiplicative inverse: $\frac{1}{a+ib} = \frac{a-ib}{a^2+b^2}$ (Realify the denominator).

2.4.4 nth roots of 1

Divide the unit circle into n sectors: Integer powers of $w = e^{2i\pi/n}$. $w^{-1} = w^{n-1} = e^{-2i\pi/n}$. As $1 + w + \cdots + w^{n-1} = w(1 + w + \cdots + w^{n-1})$, $(1 + w + \cdots + w^{n-1}) = 0$: Also for any w^x. $(a + ib)^n = re^{ni\theta}$ and $(a - ib)^n = re^{-ni\theta}$ are complex conjugates like $(a + ib)$ and $(a - ib)$.

2.5 Quaternions

Can't turn $R^{n>2}$ into a field [**Check**]. In quarternions, you loose commutativity. In Octanions, you loose associativity.

Part II

Real valued Functions and their properties

3 Properties

Because of the nature of their range, real valued functions can be characterized using some special features.
Also consider properties of functions over ordered semigroups described elsewhere.

3.1 Topological properties

Limits, continuity, smoothness, steepness. See topology ref.

3.1.1 Limits

See topology ref. Left and right handed limits. $\pm\infty$ as limits.
Limits of sums, products, quotients. Squeeze or pinching theorem: If $f(x) \leq g(x) \leq h(x)$, and if $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = b$, $\lim_{x \to a} g(x) = b$.

3.1.1.1   f:F to F: Find limits

Try Substitution, factorization, rationalization.

L'Hopital rule: If $lt_{x \to c} f(x) = lt_{x \to c} g(x) = 0, L = \lim_{x \to c} \frac{f(x)}{g(x)} = \frac{f(x)'}{g(x)'}$: from definitions or from generalized mean value thm: define f(c) = g(c) = 0, so derivatives exist around c; as $\exists L, g'(x) \neq 0$.

3.1.1.2   Closed functional f

All sublevel sets of f are closed. Equivalently, the epigraph is closed. [**Find proof**]

Consider/ visualize the epigraph: If f is continuous, dom(f) is closed, then f is closed. Also, if f is continuous, dom(f) is open, f is closed iff f converges to ∞ along every sequence converging to bd(dom(f)). [**Find proof**]

3.1.2   Continuity

See topology ref.

If f, g are continuous, then f+g, fg, f/g are continous.

3.1.2.1   Absolute continuity of f:Rm to Rn

More powerful/ specialized than uniform continuity.

$\forall \epsilon \geq 0, \exists d \geq 0 : \forall$ finite sequence of pairwise disjoint sub-intervals (x_k, y_k) : $\sum_k |y_k - x_k| < \delta \implies \sum_k |f(y_k) - f(x_k)| < \epsilon$. This can be extended to $f : R^m \to X$ for any topological space X.

3.1.2.2   Simple discontinuity

Upper and lower limits exist, but different: $\lfloor x \rfloor$. Non-simple disc: f(x): 1 if $x \in Q$, 0 else.

Fixing discontinuities. $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

3.1.2.3   Extreme value existence, boundedness

(Weierstrass) If real valued f is continuous over compact (closed and bounded in R) $S = [x_1, x_2]$, it attains maximum and minimum value somewhere in S.

[**Proof**]: As S compact, f(S) compact [See topology ref.]. So f closed and bounded. By LUB property of R, supf= M; take d_n: $M - n^{-1} \leq f(d_n) \leq M$; so $f(d_n) \to M$; by Bolzano Weierstrass take convergent subseq $(d_{n_k}) \to d$; $d \in S$ as S closed; as f cont, so $f(d) = M$. □

If S not compact, there can be: unbounded but cont f: $S = (0, 1), f(x) = x^{-1}$; cont f without max: f(x) = x on (0,1); cont but not uniformly cont: $S = (0, 1), f(x) = x^{-1}$.

3.1.2.4   Intermediate value theorem

If continuous f(x):[a, b] $\to$ R ,
$u \in [f(a), f(b)], \exists c \in [a, b] : f(c) = u$: [a,b] connected, so f([a,b]) connected.

4   Definite-Integral function

4.1   Definitions

For a fixed $f : R \to R$, the definite integral function is $R^2 \to R$. In general, it is defined for any f over a measurable space.
Simply see the definition of min-cover integrals defined elsewhere: those form superior notions of integration than what follows.

4.1.1   Integral as least Upper sum, greatest lower sum

Aka Reimann Integral.

4.1.1.1   Partitions of [a,b]

Any such partition P is specified by a set of points $a = x_0 \leq x_1..x_n = b; \Delta x_i = x_i - x_{i-1}$.

Refinement P' is a refinement of P if $P \subseteq P'$. Common refinement of $P_1, P_2 : P_1 \cup P_2$.

4.1.1.2   Upper, lower sums

Take bounded function
$f : [a, b] \to R; M_i = \sup \{f(x)|x \in [x_{i-1}, x_i]\} ; m_i = \inf \{..\}$; visualize; upper, lower sums $U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i; L(P, f) = \sum_{i=1}^{n} m_i \Delta x_i$;
$\overline{\int_a^b} f(x)dx = \inf_P \{U(P, f)\}, \underline{\int_a^b} f(x)dx = \sup_P \{L(P, f)\}$: both exist as both L and U sums are bounded and real.

4.1.1.3   Integrability

If $L = U$, $f \in \mathcal{R}$, that is: f is Reimann integrable.

4.1.1.4   Limitations

Not Reimann integrable: $I_Q(x)$ (Indicator function for rational numbers): $U = 1$, while $L = 0$ as every partition will contain a rational and an irrational number. However, this is box integrable: the corresponding box Integral is 0. It is but a special case of superior notions such as the Stieltjes integral and the min-cover integral.

4.1.2   Min-cover integral of measurable functions

Aka Lebesgue integral, box integral. Take a real valued function f over measure space (X, S, m). Suppose $f : X \to R$ is a measurable function from $(X, S, m) \to (R, S_r, m_r)$, where S_r is the sigma algebra of open sets in R, and m_r is the common box measure on R.

For $f(x) \geq 0 \forall x$, for any $E \in S$, $\int_E f(x) dm$ is the min-cover measure of portion of the space E_f in $(X \times R)$ bounded by E and $f(x)$. This measure was described for the general case in the product measure portion of the algebra survey. There is the additional restriction that the each $B_i = (B_{iX} \times B_{iR}) \in (S \times S_r)$ used to cover $E_i \subseteq E_f$ is such that for any $(x, f(x)) \in E_i$, its measure along R, $m_r(B_{iR}) \geq m_r(f(x))$: that is, they can be visualized as vertical boxes - with R being the vertical axis.

This can be extended to any function f. Let $f^+ = 2^{-1}(|f(x)| + f(x))$ and $f^- = 2^{-1}(|f(x)| - f(x))$ be the positive and negative parts of f, so that $f = f^+ - f^-$. Then, $\int_E f(x) dm = \int_E f^+(x) dm - \int_E f^-(x) dm$.

4.1.2.1   Importance

It is superior to Reimann integral. The indicator function over rationals $I_Q(x)$ is not reimann integrable, but it is box integrable as the measure $m(Q) = 0$.

4.1.3   Min-cover integral wrt to non decreasing g

Aka Lebesgue-Stieltjes integral or Radon integral.
This is defined as the min-cover integral obtained using a measure m_g which corresponds to g.

4.1.3.1   Upper and lower sums view

Aka Stieltjes integral. Take $g : [a, b] \to R$ nondecreasing. $\Delta g_i = g(x_i) - g(x_{i-1})$.
f bounded. Define $U(P, f, g) = \sum_{i=1}^{n} M_i \Delta g_i$, L(P, f, g), $\int_a^b f(x) dg$, $\underline{\int}$.
If $L = U$, $f \in R'(g)$: Stieltjes integrable.
For $\int$ to exist, f and g must not share any points of discontinuity: See how g is used to deal with discontinuity in f.

4.1.3.2   Advantages

Suppose f is $[x \in Q]$ in a certain interval X, $f \in R(g)$ where g(x) = constant over X.
If f bounded on $[a, b]$, has finite discontinuities, g conts at these points, then $f \in R(g)$: take partition with δg_i small in these pts, where $|m_i - M_i|$ maybe big but bounded; thence show $U - L < \epsilon$.

Series as Stieltjes integral Derive g from unit step functions: $g(x) = \sum_{n=1}^{\infty} c_n I(x - s_n)$, with $c_n \geq 0, \sum c_n$ convergent. Then $\int_a^b f(x) dg = \sum c_n f(s_n)$: From $\int f d(g_1 + g_2) = \int f d(g_1) + \int f d(g_2)$. $\int_a^b 1 dg = \sum c_n$.

4.2 Importance

For a fixed f, m, the integral can be viewed as a signed measure over X. This measure is significant as it captures and extends the notion of area/ volume etc.; and as described in another chapter, it happens to be closely related to the gradient of the function.

4.3 Basic properties

4.3.1 Integral sandwiched between L and U

If $P \subseteq P'$, $L(P, f, g) \leq L(P', f, g), U(P', f, g) \leq U(P, f, g)$: see geometrically. So, $\int_a^b f(x) dg \leq \overline{\int}$: take common refinement of of $P_1, P_2 : P'$, see $L(P_1, f, g) \leq L(P', f, g) \leq U(P', f, g) \leq U(P_2, f, g)$.
$f \in R(g) \equiv \forall \epsilon : \exists P : U(P, f, g) - L(P, f, g) < \epsilon. \leftarrow$ Pf: $L \leq \underline{\int} \leq \overline{\int} \leq U. \rightarrow$
Pf: As $\underline{\int} = \overline{\int}$, take $L(P_1), U(P_2)$ close to these, take common refinement.

4.3.1.1 Closeness to any sum between L and U

So, if $s_i, t_i \in [x_{i-1}, x_i], \sum |f(s_i) - f(t_i)| \Delta g_i \leq \epsilon$. Also, $|\sum f(t_i) \Delta g_i - \int_a^b f dg| < \epsilon$: both are bounded between L and U.

4.3.2 Continuity implies integrability

If f continuous on $[a, b]$, $f \in R(g)$ on $[a, b]$.
[**Proof**]: As $[a, b]$ compact, f uniformly cont; so $\forall \eta \exists d : |x - t| \leq d \implies |f(x) - f(t)| < \eta$; take $\eta : (g(b) - g(a))\eta \leq \epsilon$; take any partition P with $\Delta x_i \leq d$; so $U(P, f, g) - L(P, f, g) = \sum(M_i - m_i)\Delta g_i \leq \epsilon$. $\square$

4.3.3 Relationship with Uniform convergence

Take f_n on $[a, b]$, $f_n \to f$ uniformly.
If $f_n \in R, f \in R$: $lt_{n\to\infty} \int f_n dx = \int f dx$ as $\int (f_n - \epsilon_n) dx \leq \underline{\int} f dx \leq \overline{\int} f dx \leq \int (f_n + \epsilon_n) dx$.
So, if $\sum f_n \to g$ uniformly, $\sum \int f_n dx = \int g dx$.

4.3.4 Continuity of semi-definite integral

If $f \in \mathcal{R}$ on $[a, b]$: $F(x) = \int_a^x f(t) dt$ is continuous: f is bounded; so $F(y) - F(x) \leq M(y - x) < \epsilon$ as $y \to x$.

4.4 Connection to derivative

4.4.1 Differential of integral over f

If f is continuous at c: : $F(x) = \int_a^x f(t)dt$ is differentiable at c, $F'(c) = f(c)$.
[**Proof**]: Continuity of f states that $t - c < d \implies |f(t) - f(c)| < \epsilon$. So $|\frac{F(t)-F(c)}{t-c} - f(c)| = |\frac{1}{t-c}||\int_c^t f(u)du - f(c)(t-c)| = |\frac{1}{t-c}||\int_c^t (f(u)-f(c))du| \le \epsilon$.
□

4.4.2 Definite integral = difference in antiderivative

4.4.2.1 Anti-derivative

If $f \in R$, and if $\exists F$ differentiable on $[a,b]$ with $F' = f$, then F is called the antiderivative or the indefinite integral of f.
For any f, if there is an antiderivative F, $\{F(x) + k | k \in R\}$ is also an antiderivative. ***So, there are either 0 or ∞ antiderivatives.***

4.4.2.2 Fundamental theorem of calculus

$\int_a^b f(x)dx = F(b) - F(a)$.
[**Proof**]: Pick a partition $P = (x_i)$ with $U - L < \epsilon$. Apply Mean Value Thm to get: $F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i$; add all such terms to get $F(b) - F(a) = \sum f(t_i)\Delta x_i = \int_a^b f(x)dx + \epsilon$ by a property seen under Stieltjes Integral. □

4.4.3 Inter-measure Derivative

Aka Radon-Nikodym derivative. The min-cover integral may be used to define a derivative which connects two measures on the same measurable space: this is described in the chapter on measures in the Algebra survey.

4.5 Integration: evaluating definite integrals

Methods relevant (eg: transformation to polar coordinates) only for evaluating definite integrals of functions over product spaces are described elsewhere.
Also see section on important integrals for examples of application of these techniques.

4.5.1 Integration by parts

As definite integral equals the difference in the anti-derivative, $\int_a^b f(x)H(x)dx + \int_a^b F(x)h(x)dx =$
$\int_a^b f(x)h(x)dx = F(b)H(b) - F(a)H(a)$.
Thus, $\int_{[a,b]} udv = uv]_{[a,b]} - \int_{[a,b]} vdu$ is a useful rule when RHS is easier to integrate.

4.5.2 Sum, product, decomposition rules

Due to definite integral being equal to difference in the antiderivative: $\int_a^b f(x) +$
$h(x)dg = \int_a^b f dg + \int_a^b h dg.$ $\int_a^b f dg = \int_a^c f dg + \int_c^b f dg.$
$\int f d(g_1 + g_2) = \int f d(g_1) + \int f d(g_2).$

4.5.3 Transform the integral

Suppose you want to evaluate the definite integral I. Often, evaluating kI or
I^2 (as in the case of the Gaussian integral) is simpler. Thence one can deduce
the value of I.

4.5.4 Change of variables

4.5.4.1 Function composition: integrability

If $f \in R(g), f(x) \in [m, M]$, h continuous over [m, M], then $j(x) = h(f(x)) \in$
$R(g)$.
[**Proof**]: Take $\forall \eta \exists P : U(P, f, g) - L(P, f, g) = \sum (M_i - m_i)\Delta g_i = \delta(g(b) -$
$g(a)) \leq \eta$. (Supposing $M_i - m_i \leq \Delta$.)
h uniformly continuous over $[m, M]$, so $(h(M_i) - h(m_i)) \leq \epsilon$ for some ϵ.
Hence, $\forall \eta' \exists P : U(P, j, g) - L(P, j, g) = \sum (M_i^* - m_i^*)\Delta g_i = \epsilon(g(b) - g(a)) \leq$
$\frac{\eta \epsilon}{\delta} = \eta'$. $\square$
So, if $f, g \in R(g) : fg = 4^{-1}((f + g)^2 - (f - g)^2), |f(x)| \in R(g)$. Also,
$\int f(x)dg \leq \int |f(x)|dg$.

4.5.4.2 Integration

Take strictly increasing onto $h : [A, B] \to [a, b]$. Suppose that $f \in R$ on $[a, b]$.
Then $g = f \circ h \in R(h)$ on $[A, B]$. Then $\int_{x \in [A,B]} g(x)dh = \int_{t \in [h(A),h(B)]} f(t)dh =$
$\int_{x \in [A,B]} f(h(x))h'(x)dx$.
Core Idea: Take a partition P with upper and lower sums L, U. Use mean
value theorem to rewrite in terms of x. $\|$

Utility Change of variables is in general a extremely useful algebraic trick
- so is this particular instance. Suppose $x = h(y)$. With this, one can write:
$\int_{[A,B]} f(x)dx = \int_{y \in [h^{-1}(A),h^{-1}(B)]} f(h(y))\frac{dx}{dy}dy$, as the RHS may be easier to
evaluate.

4.5.5 Reversing limits

By convention $\int_a^b f(x)dx = -\int_b^a f(x)dx$.

4.6 Over product measure spaces

Consider $f(x) \geq 0 \forall x$. (The integral for the general case may thence be derived as earlier.)

Let (X, S, m) be the product measure space of $\{(X_i, S_i, m_i) : i \in 1, 2\}$. Let $E \in S = S_1 \times S_2$: Note that this includes only rectangles in R^2, but not circles etc..

$E_1 = \{x_1 \in X_1 : \exists x = (x_1, x_2) \in E\}$. Let $E_2(x_1)$ be similarly defined.

Then, from the properties of the product measure: (Fubini) $\int_E f(x) dm = \int_{E_1} g(x_1) dm_1$, where $g(x_1) = \int_{x_2 \in E_2} f(x_1, x_2) d(m_2)$.

[**Proof**]: Let R be associated with the usual measure space (R, S_r, m_r). (Core intuition) The min cover measure of the space in $(X \times R)$ bounded by E and f is $inf_{\{B_i\}}(m \times m_r)(\cup_i B_i)$, where $B_i = (B_{iX_1} \times B_{iX_2} \times B_{iR}) \in (S \times S_r)$ cover E. Because we are dealing with the product measure, $m_1 m_2 m_r(B_i) = m_1(B_{iX_1}) m_2 m_1(B_{iX_2} \times B_{iR})$. $\square$

Thus, integration over the product measure space reduces to integrating over one variable at a time; and the order in which these integrals are taken does not matter in this case. This case is aka multiple integration.

4.6.1 Bounds

If $f \leq h, \int f dg \leq \int h dg$. If $|f(x)| < M \implies |\int_a^b f dg| \leq M(g(b) - g(a))$.

4.6.2 Use polar coordinates

Consider $\int_{a,b} f(x, y) dx dy$, one can find an equivalent expression in polar coordinates using $x = r \sin \theta, y = r \cos \theta$ and $dx dy = dr(r d\theta)$. This is sometimes simpler to solve.

As in the case of the Gaussian integral, the form $\int_{a,b} f(x, y) dx dy$ may be derived from considering I^2.

4.7 Important integrals

4.7.1 Normal integral

Aka Gaussian integral. $\int_{-\infty}^{\infty} e^{-x^2} dx = I = \sqrt{\pi}$:

[**Proof**]: $I^2 = \int_{y=-\infty}^{y=\infty} \int_{x=-\infty}^{x=\infty} e^{-x^2-y^2} dx dy$; thence transform to polar coordinates and solve. $\square$

4.7.1.1 Importance

The Gaussian integral itself is important due to this nice integrability and the nice properties observed in the Gaussian distribution (very suitable for modeling). Further, it is useful because realated integrals can be used to smooth functions for various purposes (eg: optimization).

In calculating similar integrals, one often uses similar techniques and results about exponential and gamma probability densities.

Part III

R to R functions

5   Superclasses

Properties of the more general class of real valued functions is described in another part.
Properties of the more general class of functionals over vector spaces are described in the vector spaces survey.

6   Functions over N: Sequences, series, convergence

6.1   Convergent sequences $(s_n), (t_n)$ in a field

Also see Topology ref, properties noted in lin alg ref.
Sums, products, quotients (sans 0 denominator) of convergent sequences are convergent.

6.1.1   Boundedness

Upper and lower bounds.

6.1.2   Montonicity

Monotone increasing and decreasing sequences.
A monotone increasing sequence that is bounded above must converge: imagine points on real line; similarly monotone decreasing sequence that is bounded below.
(Bolzano, Weierstrass): Every bounded sequence has a subsequence which converges: let M be the bound; either [-M, 0] or [0, M] has ∞ elements; so use this repeatedly to find monotone, bounded subsequence; this converges. Also holds for R^n.

6.1.3   Upper and lower limits

Take set of subsequential limits E. Upper limit of s_n: $s^* = \limsup_{n \to \infty} s_n :=$ $\sup E$; similarly lower limit: $s_* = \liminf_{n \to \infty} s_n$.
If $s_n \to s, s^* = s_* = s$. E is closed, so $s_*, s^* \in E$.

6.2 Generating function of a sequence

Take sequence (a_i). Get generating function: $G(x; a_n) = \sum_{n=0}^{\infty} a_n x^n$.

6.3 Series

∞ series. A sequence of partial sums: $s_n = \sum_{i=1}^{n} a_i$.

6.3.1 Convergence

Many ideas from convergence of sequences.

6.3.2 Convergence tests

s_n converges $\equiv$ it is cauchy. Aka completeness property of R or C. Cauchy criteria for convergence of s_n: $\forall \epsilon, \exists N : |s_N - s_{N+n}| < \epsilon$. So, $|a_n| \to 0$.

Sum or product of absolutely convergent series is absolutely convergent: limit of new series is sum or product of limits (Cauchy product).

Comparison test: If $|a_n| \le c_n$: if $\sum c_n$ converges, $\sum a_n$ converges; if $\sum |a_n|$ diverges, $\sum c_n$ diverges: using Cauchy criterion.

If $a_k \ge a_{k+1} \ge 0, s_n = \sum a_n$ converges $\equiv t_n = \sum_{k=0}^{n} 2^k a_{2^k}$ converges: s_n monotonically increasing, so check boundedness: $\frac{t_k}{2} \ge s_n \le t_k$ for $n \le 2^k$.

6.3.2.1 Ratio test

Converges if $\limsup |\frac{a_{n+1}}{a_n}| < b < 1$: bound by geometric series $b^k a_n$; diverges if $> 1 : |a_n| \not\to 0$.

6.3.2.2 Root test

Take $x = \limsup |a_n|^{1/n}$: So $|a_n| \le x^n$. So, convergence if $x < 1$, divergence if $x > 1$; No info otherwise: $\sum 1/n \nearrow$, but $\sum 1/n^2 \to c$.

More powerful than ratio test:

$\liminf |\frac{a_{n+1}}{a_n}| \le \liminf |a_n|^{1/n} \le \limsup |a_n|^{1/n} \le \limsup |\frac{a_{n+1}}{a_n}|$. [**Find proof**]

6.3.2.3 Alternating series test

If $|c_i| \ge |c_{i+1}|; c_{2k+1} \ge 0, c_{2k} \le 0, \lim c_n = 0$, then $\sum c_n$ converges [**Find proof**]. So, $\frac{(-1)^n}{n}$ converges.

6.3.3 Conditional convergence

Absolute ($\sum |a_i|$) vs conditional ($\sum a_i$) convergence. Absolute convergence $\implies$ conditional convergence: ratio test.

For conditional convergent series: There are ∞ +ve and -ve numbers which converge to $+\infty$ and $-\infty$: else contradicts absence of absolute convergence or presence of conditional convergence.

6.3.4   Absolute convergence

All $a_n \geq 0$. s_n monotone: so converges $\equiv$ it is bounded.

6.3.5   Rearrangement

Take $\sum a_n$ conditionally but not absolutely convergent: so $|a_n| \to 0$, $x \leq y$. $\exists$ rearrangement of terms: $s'_n = \sum a'_k$, with $\liminf s'_n = x, \limsup s'_n = y$: take +ve $\{p_n\}$, -ve $\{q_n\}$: both diverge, collect enough p_n to go just over y; collect enough -ve numbers to go just under x; repeat: visualize as displacement between 2 points in R. So, can converge to any real number r. Eg: Rearrange terms of alternating harmonic series.
If $\sum a_n$ converges absolutely, all rearrangements converge to same limit: Any $s'_n = \sum a'_k$ is in some s_N and vice versa.

6.3.6   Series product

$\sum a_n, \sum b_n : c_n = \sum_{k=0}^{n} a_k b_{n-k}$ is the product: put all ordered sequences (a_i, b_i) on a line: the diagonals of $a_i vs b_i$ matrix are (c_n).
If $\sum |a_n|$ converges, $\sum a_n \to A, \sum b_n \to B, \sum c_n \to AB$: Take $D_n = B_n - B$, $C_n = \sum_{k=0}^{n} a_i B_{n-i} = \sum_{k=0}^{n} a_i (B + D_n) = A_n B + \gamma \to AB$.

6.4   Special series

6.4.1   Arithmatic series

Arithmatic series: $S_n = \sum_n Tn : T_n = a + nd$. Reverse the series, add respective terms, get: $2S_n = n(a + nd)$

6.4.2   Geometric-like series

$T_n = ar^n$. Multiply by r, subtract terms, get: $(1-r)S_n = a - ar^{n+1}$. Convergence as $n \to infty$: if $r < 1$. Similar techniques for $a + 2ar + 3ar^2 \ldots$.

6.4.3   Harmonic-like series

General harmonic series has: $T_n = (a + nd)^{-1}$.
$\sum 1/n$ diverges: $1 + 1/2 + 1/3 + 1/4..$: $\sum_{n=3}^{4} 1/n, \sum_{n=5}^{9} 1/n$ etc exceed 1/2.
Harmonic number: $H(n) \approx \ln n = \int x^{-1} dx$.
$\sum 1/n^p$ converges if $p > 1$, diverges if $p \leq 1$: check convergence of $\sum \frac{2^k}{2^{kp}}$.
$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$ converges $\equiv p > 1$.

6.4.4   Power series

$z \in C : \sum c_n z^n$. If this converges, got Generating Function!
Take $a = \limsup_{n \to \infty} |c_n|^{1/n}, R = 1/a$: convergence if $|z| < R$ (radius of convergence), divergence if $|z| > R$: From Root test, ratio test.

If R $= 1, c_i \geq c_{i+1} \geq 0$, then $\sum c_i z^n$ converges $\forall z$, except maybe z $= 1$. [**Find proof**]

6.4.5   Binomial theorem

For $(x + y)^n : n \in N$: from combinatorics. For $(1 + x)^n : n \in R$: Solve for coefficients in $(1 + x)^n = 1 + nx + ..$ by differentiating repeatedly.

6.4.6   e

$e = \lim_{n \to \infty} (1 + n^{-1})^n = \lim_{n \to \infty} = 1 + 1 + \frac{1}{2!}... \ 2 < e < 1 + \sum_{n=0}^{\infty} (\frac{1}{2})^n = 3$, and e is an increasing sequence: so it converges.
Observe relationship with definition using natural log in another section.

6.5   Summation tricks

6.5.1   Summation by parts

Like integration by parts. Seq $(a_n), (b_n), A_n = \sum a_n$. If $p \in [0, q], a_n = A_n - A_{n-1}, \sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$.
If (A_n) bounded seq, $b_0 \geq b_1.., \lim_{n \to \infty} b_n = 0, \sum b_n a_n$ converges: Can use summation by parts to show Cauchy criterion.

6.5.2   Using binomial summation formula

$\sum k = \sum \binom{k}{1} = \binom{k+1}{2}$. Similarly, any polynomial $\sum a_i k^i = \sum b_i \binom{k}{i}$.

6.5.3   Derivative to estimate limit of partial sums

$(1 - x)^{-1} = \sum_{i=0}^{\infty} x^i$; take derivatives to get: $(1 - x)^{-2} = \sum_{i=0}^{\infty} i x^{i-1} = \sum_{i=0}^{\infty} (i - 1) x^{i-1} - x^{i-1}$. Use similar trick to find: $i^2 x^i$.

6.5.4   Integral to estimate limit of partial sums

Also, use $\sum n \approx \int x dx$.
The series $-\log(1 - x) = \int(\frac{1}{1-x}) = \sum_{i=1}^{\infty} x^i/i$; also from McLaurin series.

7   Differential function

7.1   Definition

7.1.1   Rate of change

The differential function, aka derivative, is
$f'(x) = \frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$, if the limit exists.

7.1.2   Linear approximation view

Hence, $f(x + \Delta x) = f(x) + \Delta(x)f'(x)$, as $\Delta x \to 0$. As the primary use of the differential function is to be able to make linear/ polynomial approximations to f, one can view the differntial function as a measure of $f(x + \Delta x) - f(x)$ as $\Delta x \to 0$.

7.1.3   Other views and generalizations

See the derivatives of general functionals and functions in the vector spaces survey.

7.2   Existence: Differentiability

The the above limit exists at a certain point x, f is said to be differentiable at that point.
f can be differentiable on (a, b) but not on endpoints of $[a, b]$.

7.2.1   Relationship with continuity

If f differentiable, f continuous; but not reverse: consider $|x|$.
Also, $f = \sum_{n=0}^{\infty} g(4^n x)(3/4)^n$, where $g(x) = g(x + 2), g(x) = |x| on[-1, 1]$ is continuous but nowhere differentiable: Take $f_n = g(4^n x)(3/4)^n$; by Weierstrass M test, $\sum f_n$ converges to f; also converges uniformly; as the partial sums are continuous functions and as $\sum f_n \to f$, f is continuous.

7.2.2   Smoothness

If for all k, $D^k(x)$ exists at a certain point x, f is said to be smooth at that point.

7.3   Differential operator

7.3.1   Definition

Consider the operator D, which maps a given function f to its differential function f'. Note that the differential function may have a smaller domain than f.

7.3.1.1   Notation

So, $D(f)(x) = f'(x)$. Often, this is written as $Df(x)$.
Below, represent the scalar functions f and $D(f)$ as vectors in a vector spaces: see vector spaces and functional analysis surveys for details. Also, $f+g, fg, f/g$ etc.. are naturally defined.

7.3.1.2 Higher order differentials

These are defined by $D^k(x) = D(D^{k-1}(x)) : \forall k > 1, k \in N$.
Other notations include: $\frac{d^k f(x)}{dx^k}, f^{(k)}(x)$.

7.3.2 Inverse

Directly from the definition, $D(f^{-1}) = \frac{1}{D(f)}$.

7.3.3 Linearity

$D(f + g) = D(f) + D(g)$. This follows from the definition of the differential function.
Hence, the differential operator $D()$ is a linear map between these vector spaces. It is a matrix.

7.3.4 Other properties

If $f, g : R \to R$ differentiable at a certain point x, using : $D(fg) = fD(g) + D(f)g$. This follows from the definition of the differential function. Similarly, $D(f/g) = D(f)g - fD(g))/g^2$.

7.3.5 Composition (chain rule)

Suppose g, h are functions. $D(g(h)) = D(h)D_h(g)$, where $D_h(g)$ is the differential function of g evaluated at $h(x)$. This is proved using linear approximations $f(x + \Delta x) = f(x) + f'(x)\Delta x$ and the definition of $D(f)$.

7.3.6 Parametrically defined functions

Suppose $y = f(t)$ and $x = g(t)$. Then, $\frac{dy}{dx} = \frac{f'(t)}{g'(t)}$; as $y = f(g^{-1}(x))$.

7.4 Differentiation

Differentiation is the procedure of evaluating the differential operator for a certain function.

7.4.1 Differentiation tricks

$y = (f(x)^{g(x)})'$: take log on both sides, differentiate. Chain rule.

7.4.2 Differentials of important functions

Ab initio differentiation of sin x, cos x.

7.4.2.1   Differentiation of powers and exponents

$x^n : n \in N$ can be derived from the definition. Thus, using the expansion/ definition of e^x, $D(e^x)(x) = e^x$.
$\forall n \in R : \frac{d}{dx}x^n = \frac{d}{dx}e^{n \ln x} = nx^{n-1}$.

7.5   Other properties

7.5.1   Geometry

Geometry is described for the case of general functionals in the vector spaces survey.

7.5.2   Connection with extrema

If f differentiable, f' is 0 here (minimum / maximum singularity). See optimization ref.

7.5.3   Effect of uniform convergence

Take f_n on $[a, b]$, $f_n \to f$ uniformly.
$f_n = \sin nx/\sqrt{n} \to 0$, but $f'(x) \nrightarrow 0$.
If f'_n continuous, uniformly convergent, and if $\exists c \in (a, b) : f_n(c)$ converges (This condition ensures ye're not adding constants to f_n to keep f'_n convergent but f_n non convergent); then $\exists f : lt_{n\to\infty}f_n = f, f$ is differentiable, $\lim f'_n = f'$. Pf: Let $\lim f'_n = g$; From fundamental theorem of calculus: $\int_c^x f'_n(t)dt = f_n(x) - f_n(c)$; take limits on both sides, get: $\int_c^x g(x)dx = f(x) - f(c)$; g continuous as f'_n continuous, uniformly convergent to g; so use fundamental theorem of calculus again.

7.6   Polynomial approximation of f

7.6.1   f(x): Mean change and the gradient

7.6.1.1   Interior extremum existance

(Rolle) If f is continuous and
differentiable in $[x_1, x_2], f(x_1) = f(x_2) = 0, \exists c \in (x_1, x_2) : f'(c) = 0$.
Easy to make a visual argument.
[**Proof**]: There exists atleast one maximum and one minimum in [a, b]; if it happens to be in the interior set (a, b), $f'(x) = 0$ at this point; otherwise f is a constant function, and there is still an extremum in (a, b). $\square$

7.6.1.2   Mean value theorem

If f is continuous and differentiable in
$[x_1, x_2], \exists c \in (x_1, x_2) : f'(c) = \frac{f(x_1)-f(x_2)}{x_1-x_2}$. Easy to visualize. For proof, see

the generalization to two functions below.
Thence, linear approximation to f!

7.6.1.3 Relative to another function

If f, g continuous and differentiable: $(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x)$:
Make new function, apply Rolle. Aka Cauchy's mean value theorem.
[**Proof**]: Suppose that $f(b) = f(a) + M(g(b) - g(a))$. Now, solve for M. Take
$F(x) = f(x) - f(a) - M(g(x) - g(a))$. $F(a) = F(b) = 0$; so because of the
Interior extremum existance argument, there must exist some $c \in [a, b]$ with
$F'(c) = 0$. $\square$

7.6.1.4 Definite integral view and the mean

$\int_{[a,b]} f'(x)dx = f(b) - f(a) = f'(c)(b - a)$ for some $c \in [a, b]$. This can be
extended to integration $\int_{[a,b]} f'(x)dg(x)$ wrt another function $g(x)$ using the
mean value theorem relative to another function.
As integration can be viewed as an extension of summation, $\frac{\int_{[a,b]} f'(x)dx}{(b-a)} = f'(c)$
can be viewed as the mean of $f'(x)$.

7.6.2 Polynomial approximation

Aka Taylor theorem. $P(a) := \sum_{k=0}^{n-1} f^{(k)}(a)(b - a)^k/k!$; then $f(b) = P(a) + R$,
where $R = f^{(n)}(c)(b - a)^n/n! = \int_{[a,b]} f^{(n)}(y)(b - y)^{n-1}/(n - 1)!dy$ for some
$c \in [a, b]$.
Can then bound error term by bounding $f^{(n)}(c)$.
[**Proof**]: We want to find $f(b) - P(a)$. Note that $P(b) = f(b)$, and $P'(x) = \sum_{k \geq 0}(-f^{(k)}(x)/((k-1)!)+f^{(k+1)}(x)(b-x)^k/(k!)) = f^{(n)}(x)(b-x)^{n-1}/(n-1)!$:
note how P was cleverly defined around b rather than a to let this happen.
From mean value theorem wrt function g, we get $f(b) - P(a) = P(b) - P(a) = \frac{f^{(n)}(c)(b-c)^{n-1}/(n-1)!}{g'(c)}(g(b) - g(a))$. Then, using alternatively $g(x) = (b - x)^n$
and $g(x) = \int_{[a,x]} f^{(n)}(y)(b - y)^{n-1}dy$, we get the stated remainders. $\square$

7.6.3 Associated series

In Pf, note that, in general, c_i getting closer and closer to a as n increases; but
we cannot be sure about it: for all we know, all c_i may be very close to b.
The polynomial approximation series, aka Taylor series, is the polynomial approximation $P(x)$ as the degree $n \to \inf$.
McLaurin series: Taylor series about 0.

7.6.4 Importance

Polynomial approximation of functions described above is very important in
analyzing the solutions to many problems. This is because one can use the nth

degree approximation and upper bounds on $|f^{(n)}(x)|$ to get easy to analyze upper and lower bounds on $f(x)$.

For example, this is used to prove that solutions to certain optimization problems which arise in doing maximum likelihood estimation have desirable properties. Also, several optimization algorithms work by minimizing the polynomial (quadratic in the case of Newton's method) approximation to $f(x)$, and this analysis is naturally used there.

7.7 Extensions

Can't easily extend to general metric spaces by using a distance function $d()$: rates of change won't be negative.

7.7.1 Extension to functionals

See vector spaces ref.

8 Important functions

8.1 Classification by dependence on parameters

See probability ref.

8.2 Recurrance equations

Eg: $a'h_{i+2} + b'h_{i+1} + c'h_i = 0$ (Linear, homogenous). Get characteristic eqn by supposing $h_i = tx^i$; Get roots and multiplicity: $(r_1, 2), (r_2, 1)$; Then, $h_i = ar_1^i + bir_1^i + cr_2^i$; Solve for a, b, c with boundary conditions. Or, use telescoping sum.

8.3 Polynomial P over field K

Polynomial over field K has coefficients from K. Rational function: ratio of polynomials. For $\perp$ polynomials, see Approx theory ref. p is monic: highest degree of x has coefficient 1.

To show that P has high degree, show it has many roots.

8.3.1 Multivariate polynomials

$Q(x_1, ..x_n) = Q(x)$, let $x_{-i} =$ x sans x_i: Multivariate polynomial over field F; degree $d = \sum d(x_i)$. Have ∞ roots: $x_{-i} = r_{-i}$, Q(x) = univariate $P(x_i)$ has roots.

8.3.2 Symmetric polynomials

$Q(x_1 \ldots x_n) = Q(\pi(x_1 \ldots x_n))$: no change on permutation. Eg: Parity function P. If Q is symmetric, multilinear $\implies$ Q writable as monomial $Q'(\sum x_i)$ of same degree: intuitive b'coz vars are indistinct, (Find inductive proof).

8.3.2.1 Probability of multivariate polynomials becoming 0

(Schwartz Zippel).
$Q(x)$; $S \subseteq F$; $Pr_{r \in S}(Q(r) = 0) \leq \frac{d}{|S|}$: generalizes univariate case: $P(x_i)$ has d roots. True for d=0; assume for d=n-1; $Q(x) = \sum_k x^k Q_k(x_{-1})$; $Pr(Q(r) = 0) \leq Pr(Q(r_{-1}) = 0) + Pr(Q(r) = 0|Q(r_{-1}) \neq 0) \leq \frac{d-k}{|S|} + \frac{k}{|S|}$.

8.3.3 Roots of polynomial P over R

From group theory: P over the field Q. For $deg(P) \geq 5, \exists P : P(r) = 0, r \in R$, r can't be writ in terms of +, -, *, /, kth root. [**Find proof**]
Fundamental theorem of algebra: P of degree d has d roots in C. [**Find proof**]
If x is root of P, so is $\bar{x}$: $P(x) = \sum a_i x^i = 0 = \overline{\sum a_i x^i} = \sum \overline{a_i x^i} = \sum a_i \bar{x}^i = P(\bar{x})$. So, if deg(P) is odd, atleast 1 real root exists: also from the fact that $P = \Theta(x^n)$.
Every polynomial of odd degree $(= \Theta(x^n))$ has a real root: Consider large +ve and -ve values for x and Intermediate value theorem. Also from the fact that complex roots appear in pairs.
Tricks: The constant gives clues about the root. Use the Θ notation to consider the roots: To solve $(\frac{e}{x})^x = n^{-2}$; get $x \log \frac{x}{e} = 2 \log n$; so $x \leq 2 \log n \leq x^2$; so $x = (\log n)^{\theta(1)}$.
Geometry: loci.

8.3.3.1 Quadratic eqn

Reduce to $(ax + b)^2 = c$ and solve. So find roots of $x^3 - 1$ (Invent i). Similarly, every polynomial can be factorized to (sub)quadratics. Cubics.

8.4 Important functions over R and C

8.4.1 Extrema

$\min(a, b) = 2^{-1}(|a + b| - |a - b|)$.

8.4.2 Unit step function I(x)

$I(x) = [x > 0]$. See connection with series and Stieltjes integrals.
Integer functions: $\lfloor x \rfloor, \lceil x \rceil$.

8.4.3 Logarithm and exponential

$\ln x := \int_{t=1}^{x} t^{-1} dt$. By fund theorem of calculus, $\frac{d}{dx} \ln x = x^{-1}$. So, $\ln(ab) = \ln a + \ln b$, $\ln x^k = k \ln x$.

$e := x : \ln x = 1$. So, $\ln e^x = x$, $\frac{d}{dx} e^x = e^x$. The McLaurin series e^x.

8.4.4 Generalized binomial coefficient

$\binom{r}{k}$ for any $k \in Z, r \in R$: generalizes $\binom{r}{k}$ from combinatorics (See probability ref).

$\binom{n}{r} = \frac{n}{r} \binom{n-1}{r-1}$. For $r \in N$: $\binom{r}{k} = \binom{r}{r-k} = \frac{r}{r-k} \binom{r-1}{r-k-1} = \frac{r}{r-k} \binom{r-1}{k} \therefore r \binom{r}{k} = (r-k) \binom{r}{k}$: Both sides are degree k polynomials in r, agree in ∞ points: so should be identically equal $\forall r \in R$.

So, $\binom{r}{k} = \binom{r-1}{k} + \binom{r-1}{k-1}$. Applying repeatedly, $n \in N : \sum_{k=0}^{n} \binom{r+k}{k} = \binom{r+n+1}{n}$.

By induction: $n, m \in N : \sum_{k=0}^{n} \binom{k}{m} = \binom{n+1}{m+1}$. Useful in summing series.

$\binom{r}{m} \binom{m}{k} = \binom{r}{k} \binom{r-k}{m-k}$: combinatorial proof for r, m, k integers; thence generalize to $r \in R$: use Polynomial interpolation argument.

Vandermonde convolution: $\sum_{k} \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}$: combinatorial proof for integer case; generalize to $r \in R$: use Polynomial interpolation argument.

8.4.4.1 Polynomial interpolation argument

This is very useful in extending relationship between integers to all real numbers: as in proof of $r \binom{r}{k} = (r-k) \binom{r}{k}$.

8.4.5 Logistic sigmoid function

$f(z) = (1 + e^{-z})^{-1}$: an S trapped between 0 and 1. Used in logistic regression. $f'(z) = (1 - f(z))f(z)$.

8.4.6 Normal function

Aka Gaussian function. Generalization of Gaussian distribution.

$f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$: mean at b, max value a, variance/ width c^2. Using Gaussian integral $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

8.4.7 Gamma function

$\Gamma(z) = \int_{0}^{\infty} x^{z-1} e^{-x} dx$. $\Gamma(1) = 1$. $\Gamma(a) = (a-1)\Gamma(a-1)$: using integration by parts; so generalizes factorial: $\Gamma(n) = (n-1)!$. Useful in specifying normalization constants in some distributions.

Part IV

Bounding techniques

9 Important limits and bounds

$p^{1/n} \to 1$: take seq $x_n = p^{1/n} - 1$. Also, $n^{1/n} \to 1$.

9.1 Bounds on combinatorial quantities

9.1.1 Bounds on n!

(Stirling)$n! \approx \sqrt{2n\pi}(\frac{n}{e})^n$. [**Find proof**]

9.1.2 Bounds on (a choose b)

$\binom{a}{b} = \frac{a!}{b!(a-b)!} \approx \frac{a^a}{\sqrt{2n\pi}b^b(a-b)^{a-b}}$ for large a, b and a-b.

So: $\binom{a}{b} \leq (\frac{ea}{b})^b$. $\lceil\frac{a}{b}\rceil \leq \frac{a-1}{b} + 1$. $\frac{e^m}{\sqrt{em}} < \frac{m^m}{m!} \leq \sum \frac{m^i}{i!} = e^m$. $(\frac{a}{b})^b < \binom{a}{b} < a^b/b!$.

9.2 Bounds on powers and exponents

9.2.1 Bounds on $(1+b)^n$ and e^n

Using Taylor series of e^{-x} wrt 0, stopped at x^2: $e^{-x} \geq 1 - x$.
$(1-x^{-1})^x \leq e^{-1} \leq (1-x^{-1})^{x-1}$: first inequality from taylor series argument.
Latter ineq similarly proved for $x \in (0, 1)$ by showing $e \geq (1-x^{-1})^{1-x}$.
Also can get bounds using Taylor series; eg: $2^{-1}(e^{ac} + e^{-ac}) \leq e^{a^2c^22^{-1}}$.

9.2.2 Bounds on log x

Using log series, or from concavity: $\ln x \leq x - 1$; easy to visualize with log x curve under x - 1 line.

9.3 Inequalities among means

9.3.1 Bound ab with a convex combination of powers

Aka Young's inequality. $a, b \geq 0$; $p, q \in R^+$ with $p + q = 1$; then $ab \leq pa^{\frac{1}{p}} + qb^{q^{-1}}$. Special case of the inequality of weighted arithmetic and geometric means, but is described separately to illustrate the proof technique.
Pf: Let $a, b > 0$. $\ln ab = \ln a + \ln b = p \ln a^{p^{-1}} + q \ln b^{q^{-1}}$. $\ln(x)$ is strictly concave function on R^+; thence use this to get inequality.

9.3.2   AM GM HM inequality

The various means are described in the statistics survey.
Suppose $a, b \geq 0$. Using the same technique (taking $\log(a^p b^q)$ and taking advantage of concavity of $\log(x)$), we see that weighted GM $\leq$ weighted AM. Also, weighted GM $\geq$ weighted HM: from applying the AM-GM inequality: $(a^p b^q)^{-1} \leq pa^{-1} + qb^{-1}$.

10   Analysis/ bounding techniques

Also see functional analysis ref.
Effectively bound one function with another.

10.1   General ideas

Use convexity of functions! See proof of AM/ GM inequalities, for example.

10.1.1   Bound f over [a, b] with a constant c

Show that f is monotone over $[a, b]$, perhaps by taking derivative; take c to be value at a or b.

10.2   Asymptotics

Asymptotic notation: o, O, Θ, Ω, ω.

10.2.1   Hierarchy of growth of functions

$c, \lg^* n, \log n, n^k, n^{\log n},$
$n^{n^{1/3}} = e^{n^{1/3} \ln n}, e^n, n! = \theta(\sqrt{n}(\frac{n}{e})^n), n^n.$

10.2.2   Iterative logarithm

$\lg^* n$: 0 if $n \leq 1, 1 + \lg^* \lg n$ otherwise. Very slow growing.

Bibliography

[1] Walter Rudin. *Principles of Mathematical Analysis (International Series in Pure & Applied Mathematics)*. McGraw-Hill Publishing Co., September 1976.