Convex Optimisation - Boyd.

② CONVEX SETS.

Affine $x_1, x_2 \in A \Rightarrow x = \theta x_1 + (1-\theta)x_2 \in A$ line

Convex $0 \le \theta \le 1$. line segments.

Convex hull $\sum \theta_i = 1$. cone $x = \sum \theta_i x_i$ $\theta \ge 0$. (rays)

Hyperplane : $\{x \mid a^T x = b\}$ affine & convex

halfspace : $\{x \mid a^T x \le b\}$ + ellipsoid.

Euclidean ball : $B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\} = \{x_c + ru \mid \|u\|_2 \le 1\}$.

Norm ball : $\{x \mid \|x - x_c\| \le r\}$

Norm cone : $\{(x,t) \mid \|x\| \le t\}$ → pointed at $\underline{0}$.

Polyhedra : $Ax \le b$, $Cx = d$ $A \in R^{m \times n}$

PSD cone : S^n set of symmetric $n \times n$ matrices.

$S^n_+ = \{X \in S^n \mid X \succeq 0\}$ psd. $n \times n$ matrices (cone)

Convexity Preserving Operations

1) intersection

2) affine transform : $f: R^n \to R^m = Ax + b$

3) perspective : $P: R^{n+1} \to R^n$ $p(x,t) = \frac{x}{t}$. f - is it a project

4) Linear fractional function : $f(x) = \frac{Ax+b}{c^T x + d}$.

Generalised Inequalities

$K \subseteq R^n$ is proper if closed, solid, pointed (contains no line)

eg $K = \{x \in R^n\}$ $\sum x_i t^i \ge 0$ $\exists t \in [0,1]$

$x \le_K y := y - x \in K$ or $x <_K y \Leftrightarrow y - x \in \text{int} K$.

eg. $x \le_{R_+^n} y \Leftrightarrow x_i \le y_i$

$K = S^n_+$, $X \le_{S^n_+} Y \Leftrightarrow Y - X$ positive semi-definite for matrices Y, X

properties: $x \le_K y$, $u \le_K v \Rightarrow x + u \le_K y + v$.

$\le$ not in general a linear ordering. eg. can have $x \not\le_K y$, $y \not\le_K x$.

$x \in S$ is - minimum of S w.r.t. $\le_K$ if $\boxed{y \in S} \Rightarrow \boxed{x \le_K y}$. (S) minimum

- minimal $\boxed{y \in S \Rightarrow}\boxed{y \le_K x \Rightarrow y = x}$. - (S) minimal.

SEPARATING HYPERPLANE THEOREM, C, D disjoint convex

$a^T x \le b$, $x \in C$, $a^T x \ge b$, $x \in D$ $\{x \mid a^T x = b\}$ separates C, D.

SUPPORTING HYPERPLANE THEOREM. $\{x \mid a^T x = a^T x_0\}$.

$\{x \mid a^T x = a^T x_0\}$ and $\{a^T x \le a^T x_0 \; \forall x \in C\}$.

if C convex $\Rightarrow \exists$ supporting hyperplane $\forall x \in \partial C$.

DUAL CONES, GENERALISED INEQUALITIES.

$K^* = \{y \mid y^T x \ge 0 \quad \forall x \in K\}$

$K = \{(x,t) \mid \|x\|_2 \le t\}$, $K^* = \{(x,t) \mid \|x\|_2 \le t\}$.

MINIMUM & MINIMAL VIA DUAL CONE

? minimum w.r.t. $\le_K$ $\forall \lambda >_{K^*} 0$ x minimises $\lambda^T x$ over ?

minimal w.r.t. $\le_K$

$0 \le \theta \le 1$

③ CONVEX FUNCTIONS $f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$

convex $ax+b$, e^{ax}, x^α ($\alpha \ge 1$ or $\alpha < 0$), $|x|^p$, $x \log x$ (neg. entropy)

concave $-f$(convex) x^α on $0 \le \alpha \le 1$; $\log x$

Norms convex by $\triangle$ ineq: $\|x_1 + x_2\| \le \|x_1\| + \|x_2\|$ and they are linear $\|\theta x\| = \theta\|x\|$

Affine functions are convex and concave.

$f: R^n \to R$ convex $\iff$ $g: R \to R$ $g(t) = f(x + tv)$ convex in any t.

$f(y) \ge f(x) + \nabla f(x)^T (y - x)$ $\nabla^2 f(x) \succeq 0$ for any $x \in S^n$.

α-sublevel set: $f: R^n \to R$ $C_\alpha = \{x \in \operatorname{dom} f : f(x) \le \alpha\}$ preserve convexity

epigraph: $f: R^{n+1} \to R$ $\operatorname{epi} f = \{(x,t) \in R^{n+1} \mid x \in \operatorname{dom} f, f(x) \le t\}$.

$\boxed{f \text{ convex} \iff \operatorname{epi} f \text{ convex}}$

Jensens: $f(Ex) \le Ef(x)$

CONVEXITY CHECKS:

1. $\nabla^2 f \succeq 0$

2. ~~obtained from~~

a) $\sum w_i f(x_i)$ b) $f(Ax+b)$ c) $\max(f_i)$ d) pointwise supremum

composition e) $f(x) = h(g(x)) \Rightarrow f''(x) = h''(g(x))\, g'(x)^2 + h'(g(x))\, g''(x)$.

f) vector $f''(x) = g'(x)^T \nabla^2 h(g(x))\, g'(x) + \nabla h(g(x))^T g''(x)$

perspective $g(x,t) = t f(x/t)$ $\operatorname{dom} g = \{(x,t) \mid x/t \in \operatorname{dom} f, t > 0\}$

convex if f convex

Conjugate $f^*(y) = \sup_{x \in \operatorname{dom} f} (y^T x - f(x))$ - always convex.

quasiconvex functions: $f: R^n \to R$ if $\operatorname{dom} f$ is convex and sublevel sets $S_\alpha = \{x \in \operatorname{dom} f : f(x) \le \alpha\}$ convex $\forall \alpha$.

$\mathrm{IRR}(x) = \inf\{r \ge 0 : \sum_{i=0}^n (1+r)^{-i} x_i = 0\}$

there are many solutions.

④ CONVEX OPTIMISATION min $f(x)$ s.t. $f_i(x) \le 0$, $h_i(x) = 0$.

standard form: min $f_0(x)$ st $f_i(x) \le 0$, $a_i^T x = b_i$ f_i convex, h_i affine

quasiconvex: f_0 quasiconvex, f_i convex

feasible set convex $\Rightarrow \nabla f_0(x)^T (y - x) \ge 0$ $\forall y$ feasible.

Equivalent problems: convexity preserving transformations $\min_x f(x)$ $f_i(x) \le 0$, $Ax = b$.

① eliminate $Ax = b \Rightarrow x = Fz + x_0$

② equality constraints ③ introduce slacks for $\le$

④ epigraph form $\min_{x,t} t$ st $f_0(x) - t \le 0$

⑤ min over some variables $\min_{x_1,x_2} f(x_1,x_2) = \min_{x_1} \inf_{x_2} f_0(x_1,x_2)$ $f(x) \le 0$

LP min $c^T x + d$ st $Gx \le h$, $Ax = b$.

convex with affine objective and constraints

diet problem linear fractional

QP

QCQP $\to$ SOCP e.g. stochastic programming.

$\text{dual}(v,\lambda) = \inf_x \text{Lagrangian}$
$= g(v,\lambda) \le p^*$.

Lagrangian = constraint penalized cost function.
$h_i(x) = 0 \equiv$ null space.
$f_i(x)$ always ≤ 0.

DUALITY

$\min f_0(x)$ $\quad f_i(x) \le 0, \quad h_i(x) = 0 \quad x \in D \subset \mathbb{R}^n$

Lagrangian $\boxed{L(x,\lambda,v) = f_0(x) + \sum_i^m \lambda_i f_i(x) + \sum_i^p v_i h_i(x)}$

Lagrangian dual $g(\lambda, v) = \inf_{x \in D} L(x,\lambda,v)$ = f^n(Lagrangian multipliers = shadow costs)

lower bound $\lambda \ge 0 \quad x \in D \quad \boxed{f_0(\tilde{x}) \ge L(\tilde{x},\lambda,v) \ge \inf_{x \in D} L(x,\lambda,v) = g(\lambda,v)}$

$p^* \ge g(\lambda,v)$

EXAMPLES.

(1) $\min x'x : Ax = b \Rightarrow L(x,v) =$

Lagrangian $L(x,v) = x'x + v'(Ax-b) \quad \frac{\partial L}{\partial x} = 0 \Rightarrow x^* = -\frac{1}{2}A'v$ optimal x

L dual $\quad g(\lambda,v) = L(x = x^* = -\frac{1}{2}A'v, v) = \underbrace{-\frac{1}{4}v'AA'v - b'v}_{\text{concave}} \le p^*$ $\quad \cup f(x)$ $\quad \cap g(\lambda,v)$

(2) $\min c'x \quad : Ax = b, \quad x \ge 0$

Lagrangian $L(x,\lambda,v) = c'x + v'(Ax-b) - \lambda'x$

L is affine in $x \Rightarrow g(\lambda,v) = \inf_x L(x,\lambda,v) = \begin{cases} -b'v & \text{if } A'v - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$ ✓

$g(\lambda,v)$ linear on $A'v - \lambda + c = 0 \Rightarrow$ concave.

$\boxed{p^* \ge g(\lambda,v) = -b'v \quad \text{if } A'v + c \ge 0}$ dual LP.

(3) $\min \|x\|$ s.t. $Ax = b$

Lagrangian dual : $g(v) = \inf_x (\|x\| - v'Ax + b'v) = \begin{cases} b'v & \text{on } \|A'v\|_* \le 1 \\ -\infty \end{cases}$

where $\|v\|_* = \sup_{\|u\| \le 1} u'v$ is dual norm of $\|*\|$.

(4) $\min x'Wx \quad \forall i \; x_i^2 = 1$ non convex D has 2^n points.

W_{ij} is cost of assigning to different sets.

L dual : $g(v) = \inf_x (x'Wx + \sum_i v_i(x_i^2 - 1))$
$= \inf_x x'(W + \text{diag}(v))x - 1'v$
$= \begin{cases} -1'x & (W + \text{diag}(v)) \ge 0 \text{ is } \underline{\text{psd}} \\ -\infty & \text{otherwise} \end{cases}$

Lagrangian dual and conjugate function $\quad \min f_0(x) : Ax \le b \quad Cx = d$

$g(\lambda,v) = \inf_{x \in \text{dom}(f_0)} (f_0(x) + (A'\lambda + C'v)'x - b'\lambda - d'v)$
$= -f_0^*(-A'\lambda - C'v) - b'\lambda - d'v$ ie the infimum ie $x = -(A'\lambda + C'v)$ by inspection

* $f^*(y) = \sup_{x \in \text{dom} f} (y'x - f(x))$ CONJUGATE
$\Rightarrow$ simplifies derivation of dual

THE DUAL PROBLEM $\max g(\lambda,v)$ s.t. $\lambda \ge 0$. $\quad g(\lambda,v) = \inf_x L(x,\lambda,v)$
ie $\max_{\lambda,v} \left(\min_x (f_0(x) + \lambda' f(x) + v'h(x)) \right)$ s.t. $\lambda \ge 0$.

WEAK DUALITY $\wedge$: $d^* \le p^*$

STRONG DUALITY $\quad d^* = p^*$ - usually holds for convex problems.

* Slaters constraint qualifications

APPROXIMATION AND FITTING

Norm approximation - $\min \|Ax - b\| : A \in \mathbb{R}^{m \times n} \quad m \ge n$ - fitting etc.

Chebyshev $\quad \min 1'y$ s.t. $-y \le Ax - b \le y$.

Least-norm problems $\quad \min \|x\| : Ax = b \quad m \le n$.

can use convex penalty $\|\phi(Ax - b)\|$

Regularized approximation, robust approx.

min $x'Ax$ st $c'x=1$ A is correlation,
$L(x,\lambda) = x'Ax + \lambda(c'x$

(10) UNCONSTRAINED MINIMIZATION.

initially $x^0 \in \text{dom} f$ $S = \{x \mid f(x) \le f(x^0)\}$ is closed.

Strong convexity $\exists m$ $\nabla^2 f(x) \geq mI$ $\forall x \in S$.

$\Rightarrow f(y) \geq f(x) + \nabla f(y-x) + \frac{m}{2}\|x-y\|^2$.

step by gradient descent $\Delta x_{gd} = -\nabla f(x)$

steepest descent $\Delta x_{nsd} = \underset{v}{\text{argmin}} (\nabla f'(x) v : \|v\|=1)$

$\Delta x_{sd} = \|\nabla f(x)\|_* \Delta x$

newton $\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x)$ from

from minimise $f(x+d) = f(x) + \nabla f' d + \frac{1}{2} d' \nabla^2 f d$

implementation $\nabla^2 f \Delta x = -\Delta f$ via Cholesky.

LP

$f = c'x$ $(B \ N)\begin{pmatrix} x_b \\ x_n \end{pmatrix} = b$

$= c_b'(B^{-1}b - B^{-1}Nx_n) - c_n' x_n$ in null space

$= c_b' B^{-1} b - (c_b' B^{-1} N - c_n') x_n$

$= y' b + (c_n' - y'N)$

min $c'x$: $Ax = y$, $x \geq 0$ $\equiv$ max $y'z$: $Az \leq c$ z free.

KKT min $f(x)$: $g_i(x) \geq 0$ $L(x,\lambda) = f(x) - \sum_i^m \lambda_i g_i(x)$

$L(x,\lambda) = f(x) - \sum \lambda_i g_i(x)$

1st order necessary conditions:

(1) Dual feasible : $\lambda_i^* \geq 0$ or $f_i(x^*) \leq 0$ — dual feasible: lower bound of primal problem.

(2) Primal feasible : $g_i(x^*) \geq 0$

(3) Complementary slackness : $\lambda_i^* g_i(x^*) = 0$

(4) $\nabla_x L(x^*, \lambda^*) = 0$ ie. $\nabla f(x^*) = \sum^m \lambda_i^* \nabla g_i(x^*)$ — Jacobian of binding constraints full rank.

min $f(x)$: $g(x) = 0$ $L(x,\lambda) = f(x) + \lambda' g(x)$ wnt $\nabla f = 0$.

$\exists \lambda_0$: $\nabla_x L(x_0, \lambda_0) = 0$.

and $\forall y$ $(\nabla_x g(x))' y = 0$

$y' \nabla_x^2 L(x, \lambda_0) y \geq 0$.

ie. move along constraint until tangency $f(x) \Downarrow$ until tangent to last constraint.

* SEARCHING SIMPLEX REPLACED BY TANGENCY.

1) complimentary slackness $\Rightarrow f_0(x) = L(\bar{x}, \bar{\lambda}, \bar{v})$

2) $\nabla_x L = 0$ and convex $g(\bar{\lambda}, \bar{v}) = \inf_{\lambda v} L(\bar{x}, \lambda, v) = L(\bar{x}, \bar{\lambda}, \bar{v}) \Rightarrow f_0(\bar{x}) = g(\bar{\lambda}, \bar{v})$.

SLATER'S CQ $\Rightarrow$ x optimal $\Leftrightarrow$ $\exists \lambda, v$: KKT satisfied.

PERTURBATION AND SENSITIVITY

min $f(x)$ st. $f_i(x) \leq u_i$, $h_i(x) = v_i$, $\equiv$ max $g(\lambda, v) - u'\lambda - w'v$ st. $\lambda \geq 0$

$p^*(u,v) \geq g(\lambda^*, v^*) \geq g(\lambda^*, v^*) - u'\lambda^* - w'v^* = p^*(0,0) - u'\lambda^* - v'v^*$.

$\lambda_i^* = \partial p^*(0,0) / \partial u_i$ $v_i^* = \partial p^*(0,0) / \partial v_i$.

REFORMULATIONS FOR BETTER DUALS. $f_0(x) \Rightarrow \phi(f_0(x))$ ϕ convex awesome !

DUALITY AND ADJOINTS – adjoints of model sensitivity / optimization

linear functionals = moments = samples.

vector space $L(X, \mathbb{R})$

dual space of X, X^*

norm of X^* $\|y\|_{X^*} = \sup_{x \in X, \|x\|_X = 1} |y(x)| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|y(x)|}{\|x\|}$

$X \in \mathbb{R}^n$ $\forall v \in \mathbb{R}^n$

duality for Hilbert spaces:

Riesz-Representation $l_\psi(x) = (x, \psi)$ $x \in X, \psi \in X \Rightarrow$ dual of $X = X$!.

Riesz-representors (averages).

$\psi = \chi_\omega / |\omega|$ where χ_ω is characteristic function of ω (average)

$\psi = \delta_c = \int_c f(s) ds$

$\psi = f / \|f\|$

Bracket notation $\forall x \in X, y \in X^* \Rightarrow y(x) = \langle x, y \rangle$

$|\langle x, y \rangle| \leq \|x\|_X \|y\|_{X^*}$, $x \in X, y \in X^*$!

MOTIVATION FOR ADJOINT:

X, Y normed vector spaces $L \in L(X, Y)$ continuous linear map.

compute $l(L(x))$ for some $x \in X$.

QUESTIONS ① efficiently compute sample

② what is error in sample

③ x conditioned on sample

④ ~~given~~ $y \longrightarrow L$

X^*, Y^* dual spaces.

$\Rightarrow \exists y^* \in Y^*, x^* \in X^*$ s.t. $\boxed{x^*(x) = y^*(L(x))}$ $\langle x^*, x \rangle = \langle y^*, y \rangle$ def of adjoint

adjoint map $L^*: Y^* \to X^*$ satisfies: $\langle L(x), y^* \rangle = \langle x, L^*(y^*) \rangle$. $x \in X$, $y \in Y^*$

PROPERTIES OF ADJOINT OPERATORS

X, Y, Z are normed linear spaces $L_1, L_2 \in L(X, Y)$.

- more complicated for differential operators.
- need to assume spaces sufficiently big.

eg. Hahn-Banach theorem often used

ADJOINTS FOR DIFFERENTIAL OPERATORS. $(Lu, v) \to (u, L^* v)$

L^* obtained via integration by parts, STEPS:

① Formal adjoint assuming compact support – ignore BCs

② Compute adjoint BCs to make bilinearity hold.

e.g. KBE, KFE

EVOLUTION OPERATOR:

$\int_0^T \frac{du}{dt} u \, dt = u(t) v(t) |_0^T - \int_0^T u \frac{dv}{dt} dt$. – runs backward $v(T) = 0$

dual space ~ use: X is normed vector space over $\mathbb{R}$

$\Rightarrow X^*$ is a Banach space (even if X isn't).

CONDITION OF AN OPERATOR:

adjoint problem ~ stability of original

Theorem: singular values of L are $S = \sqrt{L^* L}$

$\boxed{\text{LINEAR PROBLEMS}}$. $L \in \mathcal{L}(X,Y)$ $Lx = b$.

$b \in \text{range}(L) \Rightarrow \boxed{y^*(b) = 0 \ \forall y^* \in \text{null}(L^*)}$

sufficient is range(L) closed in Y.

e.g. $Ax = b$ solvable $\Leftrightarrow$ ($b \perp$ to y: $A^T y = 0$) — range of $A \in$ orthogonal complement of null space of A^T

e.g. X is Hilbert space: $L \in \mathcal{L}(X,Y) \Rightarrow$ range of $L^* \subset$ (orthogonal complement of null space of L)

$\Rightarrow$ Range(L^*) large $\Rightarrow$ Null$(L^*)^\perp$ large $\Rightarrow$ Null(L^*) small.

$\Rightarrow$ existence of sufficient many ϕ st $L^*\phi = 0 \Rightarrow$

$\exists$ at most one u st $\boxed{Lu = b}$.

AUGMENTED SYSTEM. (primal-dual)

$$\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}$$

$\Rightarrow$ adjoint solvability condition drops out easy.

orig and adjoint balance each other.

+ underdetermined: $AA^t y = b$, $Ax = b$, $L(L^*(y)) = b$ $x = L^*(y)$.

for differential operators:

$\text{div} F = \rho$ $F = \text{grad}\, u \Rightarrow \text{div grad}\, u = \Delta u = -\rho$.

GREENS FUNCTIONS. $Lx = b$ linear functional $l(\cdot) = (\cdot, \psi) \Rightarrow$

$L^*\phi = \psi$. $\Rightarrow$ variational analysis

$\Rightarrow \boxed{l(x) = (x, \psi) = (x, L^*\phi) = (Lx, \phi), (b, \phi)}$.

$-\Delta u = f$, $x \in \Omega$, $u = 0$, $x \in \partial\Omega$

Green's function solves $-\Delta\phi = \delta_x$ (delta function at x).

$\Rightarrow u(x) = \boxed{(u, \delta_x) = (f, \phi)}$ — must impose adjoint BCs.

$\boxed{\text{NON-LINEAR PROBLEMS}}$. —no natural adjoint for general non-

$F(e) = f(u+e) - f(u)$ $\text{dom}\, F = \{v \in X : v + u \in f\}$.

$A^*(e)$ $(F(e), w) = (e, A^*(e) w)$. adjoint associated with F not adjoint of F.

$A = USV' = \sum s_i \langle x, v_i \rangle u_i$.

Lax: $u^{n+1} = Qu^n + \Delta t G^n$ consistent + stable $\Rightarrow$ converge[illegible]

Consistency: stability $\exists K, k$ st $\|u^{n+1}\| \leq$ [illegible]

Thomas - Parabolic PDEs $\quad e^{i\theta} = \cos\theta + i\sin\theta$

$Lv = F \Rightarrow L_h^n u_h^n = G_h^n$ $\quad$ convergent

convergent. $\|u^{n+1} - v^{n+1}\| = O(\Delta x^p) + O(\Delta t^q)$ i.e. $\exists C$ st $\|\cdot u\| \leq C(\Delta x^p, \Delta t$

$\|Q\| = \max_{\|u\|\leq 1} \|Qu\|$. $\quad \|A\| = \sigma(A) = \max\{|\lambda|: \lambda$ eigenvalue$\}$?

eigenvalues of tridiagonal matrix. $\lambda_j = b + 2c\sqrt{\frac{a}{c}} \cos\frac{j\pi}{N+1}$

consistent if $\forall \phi$ $(L\phi - F)|_k^n - (L_k^n \phi(k\Delta x, n\Delta t) - G_k^n) \to 0$ as $\Delta x, \Delta t \to 0$. $\quad$ * good approx

$v^{n+1} = Qv^n + \Delta t G^n + \Delta t \tau^n$

study out this summit[illegible]

$v_t = \nu(v_{xx} + v_{yy}) + F(x,y,t)$

$v(x,y,t) = g(x,y,t)$ on ∂R $\quad e^{-iwx} \cdot v(x,t)\,dx$

$v(x,y,0) = f(x,y)$

$$\mathcal{F}(\underline{u})(w,v) = \hat{u}(w,v) = \frac{1}{2\pi} \sum_{j,k=-\infty}^{\infty} e^{-i(jw+kv)} u_{jk}$$

$$\mathcal{F}(\underline{u})(\underline{w}) = \hat{u}(\underline{w}) = \frac{1}{2\pi} \sum_{\underline{j}=-\infty}^{\infty} e^{-i(\underline{w}\cdot\underline{j})} u_{\underline{j}}$$

$\int e^{ax}$

$d/dx\, e^{f(x)} = f'(x) e^{f(x)}$

$d/dx \int (g(x)) = f$

③ STABILITY - all functions/vectors in L_2 or l_2 (discrete).

$v_t = v_{xx} \quad x \in R \quad v(x,0) = f(x) \qquad \hat{v}(w,t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-iwx} v(x,t)\,dx$.

$\hat{v}_t(w,t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-iwx} v_{xx}(x,t)\,dx \qquad v(x,t) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{iwx} \hat{v}(w,t)\,dx$

$= (2\pi)^{-1/2} (-w^2) \int_{-\infty}^{\infty} e^{-iwx} v(x,t)\,dx$ $\quad \int x \int \frac{1}{2} x^2$

$= -w^2 \hat{v}(x,t) \qquad \hat{u}(w) \Leftrightarrow u(k)$

Parsevals $\|v\|_2 = \|\hat{v}\|_2$

Discrete FT: $\hat{u}(w) = (2\pi)^{-1/2} \sum_{n=-\infty}^{\infty} e^{-inw} u_n \quad w \in [-\pi, \pi]$. a function.

Formal proof of inversion formula: $u_k = (2\pi)^{1/2} \int_{-\pi}^{\pi} e^{ikw} \hat{u}(w)\,dw$.

$u_k = (2\pi)^{-1/2} \int_{-\pi}^{\pi} e^{ikw} \hat{u}(w)\,dw = (2\pi)^{-1/2} \int_{-\pi}^{\pi} e^{ikw} (2\pi)^{-1/2} \sum_{n=-\infty}^{\infty} e^{-inw} u_n\,dw$

$= (2\pi)^{-1} \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} e^{i(k-n)w} u_n\,dw$

$= (2\pi)^{-1} \sum_{n=-\infty}^{\infty} u_n \Big[\frac{1}{i(k-n)} e^{i(k-n)w}\Big]_{-\pi}^{\pi}$

$= (2\pi)^{-1} \sum_{n=-\infty, n\neq k}^{\infty} \frac{1}{i(k-n)} \left(e^{i(k-n)\pi} - e^{-i(k-n)\pi}\right) + u_k$

$= (2\pi)^{-1} \sum_{n=-\infty, n\neq k}^{\infty} \frac{1}{i(k-n)} \big(\cos(k-n)\pi - \cos(n-k)\pi + i(\sin(k-n)\pi - \sin(n-k)\pi)\big)$ $\quad 0$ [illegible]

$= u_k$.

Parsevals:

$\|\hat{u}\|_2^2 = \int_{-\pi}^{\pi} |\hat{u}(w)|^2\,dw$

$= \int_{-\pi}^{\pi} \overline{\hat{u}(w)}\, (2\pi)^{-1/2} \sum_{n=-\infty}^{\infty} e^{-inw} u_n\,dw$

$= (2\pi)^{-1/2} \sum_{n=-\infty}^{\infty} u_n \int_{-\pi}^{\pi} \overline{\hat{u}(w)\, e^{-inw}}\,dw$

$= \overline{\sum_{n=-\infty}^{\infty} u_n (2\pi)^{-1/2} \int_{-\pi}^{\pi} e^{inw} \hat{u}(w)\,dw}$

$= \sum_{n=-\infty}^{\infty} u_n \bar{u}_n$

$= \|u\|_2^2$

energy norm $\quad \|u\|_{2,\Delta x} = \left(\sum_{k=1}^{N} |u_k|^2 \Delta x\right)^{1/2}$

stability: $\|u^{n+1}\|_{2,\Delta x} \leq K e^{\beta(n+1)\Delta t} \|u^0\|_{2,\Delta x}$.

but $\|u\|_{2,\Delta x} = \sqrt{\Delta x}\, \|u\|_2 = \sqrt{\Delta x}\, \|\hat{u}\|_2$

$\Rightarrow$ 3.1.4: $\{u^n\}$ is stable in $l_{2,\Delta x} \Leftrightarrow \{\hat{u}^n\}$ is stable in $L_2(-\pi,\pi)$.

* no need to use inverse transform!

$z = r(\cos\theta + i\sin\theta) = x + iy = re^{i\theta}$.

series $F(x) = \frac{a_0}{2} + \sum_{j=1}^{\infty}(a_j\cos(jx) + b_j\sin(jx)) = \sum_{-\infty}^{\infty} c_j e^{ijx}$

$\ln(\cos x + i\sin x) = ix$.

$e^{ix} = 1 + ix + \frac{(ix)^2}{2!}$

$= (1 - \frac{x^2}{2!} + \frac{x^4}{4!} \dots) + i(x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots))$

$= \cos x + i\sin x$

$e^z = \lim_{n\to\infty}(1 + \frac{z}{n})^n \qquad e^{ix} = \cos x + i\sin x$.

or $d/dx\, e^{ix} = ie^{ix}$

$f(x) = (\cos x - i\sin x)e^{ix}$

$d/dx\, f(x) = (\cos x - i\sin x)ie^{ix} + (-\sin x - i\cos x)e^{ix} = 0$

$\Rightarrow f(x) = 1 = (\cos x - i\sin x)e^{ix} \Rightarrow \cos x + i\sin x = e^{ix}$.

cos

Example 3.1.1. Stability of $u_k^{n+1} = au_{k-1}^n + (1-2a)u_k^n + au_{k+1}^n \quad a = \frac{\alpha\Delta t}{\Delta x^2}$

(1) Discrete FT of both sides.

$\hat{\mathcal{F}}(u) = (2\pi)^{-1/2}\sum_{k=-\infty}^{\infty} e^{-ikw} u_k = \hat{u}(w)$.

$\hat{\mathcal{F}}(u_{\pm 1}) = (2\pi)^{-1/2}\sum_{m=-\infty}^{\infty} e^{-i(m\mp 1)w} u_m \qquad m = k\pm 1 \Rightarrow k = m\mp 1$

$= e^{\pm iw}(2\pi)^{-1/2}\sum_{m=-\infty}^{\infty} e^{-imw} u_m$

$= e^{\pm iw}\hat{u}(w)$

$\Rightarrow \hat{u}^{n+1}(w) = ae^{-iw}\hat{u}^n(w) + (1-2a)\hat{u}^n(w) + ae^{iw}\hat{u}^n(w)$.

$= (ae^{-iw} + (1-2a) + ae^{iw})\hat{u}^n(w)$

$= (2a\cos w + (1-2a))\hat{u}^n(w)$

$= (1 - 2a(1-\cos w))\hat{u}^n(w). \qquad 1-\cos w = 2\sin^2\frac{w}{2}$

$\rho(w) = 1 - 4a\sin^2\frac{w}{2}$

$\hat{u}^{n+1}(w) = (1 - 4a\sin^2\frac{w}{2})^{n+1}\hat{u}^0(w)$

And a in $|1 - 4a\sin^2\frac{w}{2}| \le 1$ ie $2 \ge 4a\sin^2\frac{w}{2}\ \forall w$ i.e. $r \le 1/2$.

$\hat{\mathcal{F}}(u) = (2\pi)^{-1/2}\sum_{n=-\infty}^{\infty} e^{-inw}u_n = \hat{u}(w)$

$\mathcal{F}(S_{\pm}u) = e^{\pm iw}\mathcal{F}(u)$.

e.g. $v_t + \beta v_x = 0 \qquad a < 0. \qquad b = \frac{\beta\Delta t}{\Delta x}$

$\Rightarrow u_k^{n+1} = (1+b)u_k^n - bu_{k+1}^n$

$\hat{\mathcal{F}}(u^{n+1}) = \hat{u}^{n+1}(w) = (1+b)\hat{u}^n - be^{iw}\hat{u}^n$

$= ((1+b) - b\cos w - ib\sin w)\hat{u}^n$.

$\Rightarrow |\rho|^2 = (1+b)^2 - 2b(1+b)\cos w + b^2$

$\frac{d|\rho|^2}{dw} = 0 \Rightarrow$ ~~maxim~~ extrema at $-\pi, 0, \pi$.

$\rho(0) = 1 \quad \rho(\pm\pi) = |1 + 2b| \Rightarrow -1 \le 1 + 2\beta \le 1$.

$-2 \le 2\beta \le 0 \quad \Rightarrow -1 \le b \le 0$.

$e^{i\omega} = \cos\omega + i\sin\omega$.

$(\ )$

$e^{i\omega} + e^{-i\omega}$
$= \cos\omega + i\sin\omega + \cos(-\omega) + i\sin(-\omega)$

explicit

Convective - diffusion. $v_t + av_x + bv_{xx} = 0$.

$\Rightarrow u_k^{n+1} = u_k^n + \frac{a\Delta t}{2\Delta x}\delta_0 u_k^n + \frac{b\Delta t}{\Delta x^2}\delta^2 u_k^n = 0$.

$\alpha/2$ $\quad \beta$ odd cancel

$\Rightarrow \hat{u}^{n+1} = \hat{u}^n + \alpha(e^{i\omega} - e^{-i\omega})\hat{u}^n + \beta(e^{i\omega} - 2 + e^{-i\omega})\hat{u}^n$

$= (1 + \alpha(e^{i\omega} - e^{-i\omega}) + \beta(e^{i\omega} - 2 + e^{-i\omega}))\hat{u}^n$

$\Rightarrow \rho(\omega) = (1 - 2\beta) + 2\cos\omega + i\beta\sin\omega)$

$\Rightarrow |\rho(\omega)|^2 = (1-2\beta)^2 +$

Implicit Scheme for $v_t + av_x + bv_{xx} = 0$.

$-\frac{a\Delta t}{\Delta x^2}\theta u_{k-1}^{n+1} + (1 + 2\frac{a\Delta t}{\Delta x^2}\theta)u_k^{n+1} - \theta\frac{a\Delta t}{\Delta x^2}u_{k+1}^{n+1} = (1-\theta)(\frac{a\Delta t}{\Delta x^2}u_{k-1}^n + \frac{a\Delta t}{\Delta x^2}u_{k+1}^n - 2\frac{a\Delta t}{\Delta x^2}u_k^n) + u_k^n$.

$\Rightarrow -\alpha\theta u_{k-1}^{n+1} + (1 + 2\alpha\theta)u_k^{n+1} - \alpha\theta u_{k+1}^{n+1}$
$= \alpha\theta' u_{k-1}^n + (1 - 2\alpha\theta')u_k^n + \alpha\theta' u_{k+1}^n$

$\hat{\mathcal{F}}$ $\Rightarrow (-\alpha e^{-i\omega}\theta + (1 + 2\alpha\theta) - \alpha e^{i\omega}\theta)u^{n+1} = (e^{-i\omega}\alpha\theta' + (1 - 2\alpha\theta') + e^{i\omega}\alpha\theta')u^n$

$\Rightarrow (1 + \theta\alpha(2 - 2\cos\omega))u^{n+1} = (1 + \theta'\alpha 2(-1 + \cos\omega))u^n$

$\Rightarrow 1 + 2\theta\alpha(1 - \cos\omega))u^{n+1} = (1 - 2\theta'\alpha(1 - \cos\omega))u^n$.

$\Rightarrow 4\alpha(1 - 2\theta) \le 2$.

$\alpha \le \frac{1}{2(1-2\theta)}$

if $\theta > 1/2 \Rightarrow$ unconditionally stable

eg. 3.1.5. $v_t = av_{xx} + cv$ $\quad \alpha = a\Delta t/\Delta x^2 \quad \gamma = c\Delta t$

$u_k^{n+1} = \alpha u_{k-1}^n + (1 - 2\alpha + c\Delta t)u_k + \alpha u_{k+1}^n$ $\quad$ 3.1.45.

$= (\alpha S_{-1} + (1 - 2\alpha + c\Delta t) + \alpha S_{+1})u_k$

$\hat{\mathcal{F}}$ $\rho(\omega) = (\alpha e^{-i\omega} + \alpha e^{i\omega} + 1 - 2\alpha + c\Delta t)$

$= (2\alpha\cos\omega + (1 - 2\alpha + c\Delta t))$ $\qquad (1 - \cos\omega) = 2(1 - \sin^2\frac{\omega}{2})$

$= (2\alpha(1 - \cos\omega) + 1 + c\Delta t)$ $\qquad = 2\cos^2\frac{\omega}{2}$

$= 1 - 4\alpha\sin^2\frac{\omega}{2} + c\Delta t$.

$\Rightarrow \alpha < \frac{1}{2}$

$\Rightarrow |\rho(\omega)| \le 1 + b\Delta t \le e^{b\Delta t}$ by expansion of $e^{b\Delta t}$.

$\Rightarrow \|\hat{u}^{n+1}\|_2 \le e^{b\Delta t}\|\hat{u}^n\|_2 \le e^{b(n+1)\Delta t}\|\hat{u}^0\|_2$.

$\Rightarrow$ 3.1.45 is stable with $K=1$, $\beta = b$

p3.1.6. $u^{n+1} = Qu^n$ difference scheme stable $\Leftrightarrow \exists\, \Delta t_0, \Delta x_0\ \exists \beta, C$

$|\rho(\omega)|^{n+1} \le Ke^{\beta(n+1)\Delta t}$

$|\rho(\omega)| \le 1 + C\Delta t$. $\leftarrow$ stronger.

* if $u^{n+1} = Qu^n$ stable $\Rightarrow u^{n+1} = (Q + b\Delta t I)u^n$ stable.

$\rho' = \rho + b\Delta t$.

$\Rightarrow |\rho'| \le |\rho| + |b|\Delta t \le 1 + C\Delta t + |b|\Delta t$

$|a + b| < |a| + |b|$.

$u = k(u-\theta)dt + \xi\, dW$

IBVP

$|\rho(\omega)| \le 1 + c\Delta t$ = von neumann stability.

$u^{n+1} = Q_1^{-1} Q u^n$.

$\sigma(Q) = \max \{|\lambda| : \lambda \text{ eigenvalue of } Q\}$ $\quad \sigma(Q) \le \|Q\|_2 = \sum d_i^2$.

if Q symmetric

$\Rightarrow$ just get the eigenvalues of a matrix.

* eigenvalues of tridigonal matrix $\quad Ax = \lambda x \quad (A - \lambda I)x = 0$

$\mathcal{F}(\underline{u})(\underline{\omega_i}) = \hat{u}(\underline{\omega}) = (2\pi)^{-n} \sum_{j=-\infty}^{\infty} e^{-i(\underline{\omega}\cdot \underline{j})} u_j$

$\rho = 1 + 2 \sum_{jm=1}^{Nm} \alpha_j (\cos \omega_j)$
$\quad = 1 - 4\left(\alpha_m \sin^2 \frac{\omega_m}{2}\right)$

$\Rightarrow \left(\alpha_x + \alpha_y \le \frac{1}{2}\right) \qquad \alpha = \frac{\alpha \Delta t}{\Delta x^2}$ - we add in different directions.

the mesh structure $\rightarrow$ time integrated and space derived coeffs.

$v_t + b\cdot \nabla v + a \nabla^2 v = 0$.

$\rho = 1 - \cancel{4} \sum_m \alpha_m \sin^2 \frac{\omega_m}{2} - i \sum_m \beta_m \sin \omega_m$

$\boxed{|\rho|^2 = \left(1 - \sum_m \alpha_m \sin^2 \frac{\omega_m}{2}\right)^2 + \sum_m \beta_m \sin \omega_m} \le 1$.

$\Rightarrow$ big constraint on α_m, β_m shrinking Δt works.

Piterbarg

ŁARNs to

MONTE CARLO

SLLN $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \xrightarrow{a.s.} \mu$ because $Var(\bar{Y}_n) = \sum_i Var\left(\frac{Y_i}{n}\right) = \sum \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n} \to 0$.

CLT $(\bar{Y}_n - \mu) \xrightarrow{d} \frac{\sigma}{\sqrt{n}} N(0,1)$ because $Var\left(\sum \frac{Y_i - \mu}{\sqrt{n}}\right) = \sum \frac{\sigma^2}{n} = \sigma^2$.

(?) why $S_n = \left(\frac{1}{n-1}\sum_i^n (Y_i - \bar{Y}_n)^2\right)^{1/2}$ — why $n-1$

std error $\bar{Y}_n - \mu \xrightarrow{d} S_n/\sqrt{n}\, N(0,1)$ ** most important.

From CLT and convergence in distribution:

u_γ s.t. $\Phi(u_\gamma) = 1 - \gamma$ $\gamma/2$ u_γ

$P\left(\bar{Y}_n \notin \left[\mu - u_{\gamma/2} \frac{S_n}{\sqrt{n}}, \mu + u_{\gamma/2} \frac{S_n}{\sqrt{n}}\right]\right) = \gamma$.

SAMPLE GENERATION

Random Number Generation

1) inverse transform $z = F^{-1}(u)$

easy to invert $F(z) = 1 - e^{-z\lambda}$ $F^{-1}(z) = \frac{1}{1-e^{-z\lambda}} = \frac{(1-e^{-z\lambda}) + e^{-z\lambda}}{1-e^{-z\lambda}} = 1 + \ldots$

Gaussian invert $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-x^2}\,dx$

(*) 2) acceptance rejection. $ce(z) \geq f(z)$ accept $e(z)$ if $u \leq \frac{f(z)}{ce(z)}$ we know the ratio

~~$E_f(g(x)) = \int f(x) g(x)\,dx$~~ accep with prob $\frac{f(x)}{ce(x)}$

~~$= \int ce(x) g(x) 1$~~ $($ $X \sim ce(z))$

$X \sim e(x)$

$p(Y \in A) = P(X \in A \mid u \leq f(X)/ce(X))$

$= P(X \in A, u \leq f(X)/ce(X)) / P(u \leq f(X)/ce(X))$

$= P(X \in A, u \leq f(X)/ce(X))$

$P(u \leq f(X)/ce(X)) = \int f(x)/ce(x)\, e(x)\,dx = 1/c$

$P(Y \in A) = P(X \in A, u \leq f(X)/e(X))$

$= c\int_A f(x)/e(x)\, e(x)\,dx$

$= \int_A f(x)\,dx$

Chi-squared $Q = \sum_i^k z_i^2$ $Q \sim \chi^2(k)$

Gamma distribution $p(\chi^2 \leq z) = \frac{1}{2^{k/2}\Gamma(k/2)}\int^{z} \ldots e \ldots dy$

→ for gamma distribution – draw from $e(z) = \lambda e^{-\lambda z}$.

3) composition

$z = \sum X_i^2$ χ^2 ... $z = -a \sum \ln u_i$

linear transform: $Z \sim N(\mu, C)$

$Y = \mu + AZ$ $E(Y) = \mu$ $Var(Y) = E((Y-\bar{Y})^2)$ — A

d-dim Z $= EY^2 - \bar{Y}^2$

p-dim p $= E((\mu + AZ)^T(\mu + AZ))$

$= E(\mu^T\mu + \mu^T AZ + (AZ)^T\mu + Z'A'AZ)$

$= E(2\mu' AZ + A'Z'ZA)$

$=$

$CC^T = \Sigma \Rightarrow Z = \mu + CX \sim N(\mu, CC^T)$.

Cholesky $-\Delta$

$$C = \begin{pmatrix} \sigma_1 & 0 \\ \sigma_2\rho & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} \sigma_1 & \sigma_2\rho \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

if $r = \text{rank}[\Sigma] < p$

$\Rightarrow Z = \mu + MY \quad M \in \mathbb{R}^{p \times r} \quad Y \sim N(0, \Sigma_Y) \in \mathbb{R}^r \quad \Sigma_Y = C_Y C_Y'$.

$\Rightarrow Z = \mu + MC_Y X \quad X \in \mathbb{R}^r$.

M from SVD

Eigenvalue decomp. $\Sigma = U \Lambda U^T$

PCA $Z \sim N(0, \Sigma) \in \mathbb{R}^p$

get closest $Z \approx DX \quad X \sim N(0, \Sigma) \in \mathbb{R}^r \quad r \leq p$.

minimize $f(D) = tr((\Sigma - DD^T)(\Sigma - DD^T)^T)$.

do this!

3.2 PATH GENERATION - INTEGRATING SDES.

$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t \quad X \in \mathbb{R}^p \quad W \in \mathbb{R}^d$

A stability etc.

3.2.3 Euler Scheme $\hat{X}_{i+1} = \hat{X}_i + \mu(\hat{X}_i)\delta t + \sigma(i\Delta, \hat{X}_i)(W_{i+1} - W_i))$.

weak convergence, strong convergence.

linear SDEs: $dX = K(\theta(t) - X(t))dt + \sigma(t, X_t)dW_t$ euler unstable for big K

$Y = Xe^{\int K dt} \Rightarrow dY_t = XKe^{\int K dt} + e^{\int K dt} dX$

$dY_t = K\theta_t e^{\int_0^t K_s ds} dt + e^{\int_0^t K_s ds}\sigma(t, X_t) dW_t$.

$Y_t = \int_0^t K_s \theta_s e^{\int_0^s K_u du} ds + \int_0^t e^{\int_0^s K_u du}\sigma(s, X_s) dW_s$.

$\Rightarrow X_t e^{\int_0^t K ds} - X_0 = \int_0^t K_s \theta_s e^{\int_0^s K_u du}$ etc.

log-euler - maintains positivity $+ O(\Delta x^2) \quad \forall \Delta x \exists K \; O(\Delta x^2) < K \Delta x^2$.

better is $\boxed{Z_t = X_t e^{\int K ds} - \int_0^t K_s \theta_s e^{\int_0^s K_u du} ds} = \int_0^t e^{\int_0^s K_u du}\sigma(s, X_s) dW_s$

simulate the mean centred at zero.

CONSISTENCY, cloud error limit.