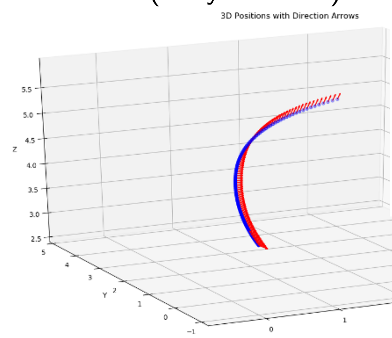
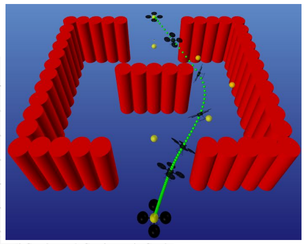
APPM 5630
"Advanced Convex Optimization"
Prof. Becker, spring 2023
Student projects

Trajectory Optimisation using Successive convexification
(Divya Tiwari)





Convex Optimization for Two-sided Fair Ranking
(Mary Monroe, Joshua Sun)

Two-sided fairness:

- Improves the worst-off users (Spotify listener)
- Improves the worst-off items (Spotify songs)

- Proposed concave welfare function, which captures user and item fairness:

$$\forall \boldsymbol{u} \in \mathbb{R}_+^n: \; W_\theta(\boldsymbol{u}) = (1-\lambda)\sum_{i\in\mathcal{N}} \psi(u_i, \alpha_1) + \lambda \sum_{j\in\mathcal{I}} \psi(u_j, \alpha_2) \text{ with } \psi(x,\alpha) = \begin{cases} x^\alpha & \text{if } \alpha > 0 \\ \log(x) & \text{if } \alpha = 0 \\ -x^\alpha & \text{if } \alpha < 0 \end{cases}.$$

Bayesian Optimization: an introduction
(Alex McManus, Leo Crowder, Jack Quinn)

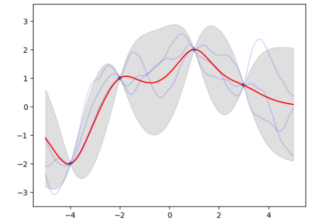


Image Classification using SVM with Linear and Nonlinear Kernel Functions
(Labib Sharrar and Ali Abbasi)

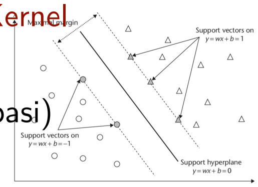
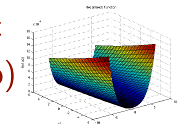


Table (1): Performance of different kernels on fitting MNIST and CIFAR datasets.

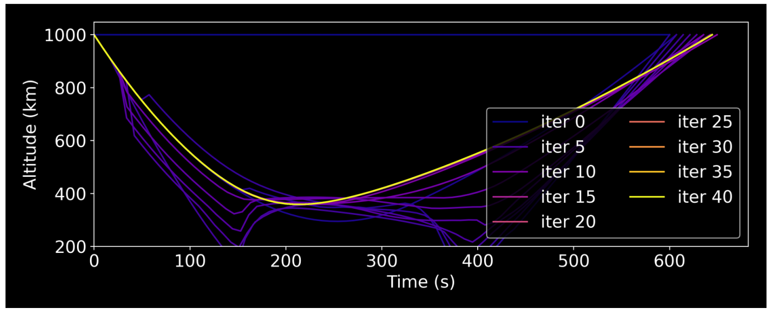
	MNIST		CIFAR	
	accuracy	Time (s)	accuracy	Time(s)
Linear	84.2%	223	85.4%	197
Radial	88.5%	450	89.9%	823
Polynomial	93.3%	345	88.6%	510
Intersection	99.1%	434	90.2%	627

Demonstrating How Exponential Moving Average Based Gradient Algorithms Fail in Certain Convex Settings (e.g. ADAM/RMSProp)
(Alex McManus, Leo Crowder, Jack Quinn)



Optimizer	2-norm of difference of the weights	∞-norm of the difference of the weights
GD	1.4430×10^{-13}	1.25011×10^{-13}
Adam	1.9033×10^{-6}	5.1294×10^{-6}
AMSGrad	8.2960	1.7580

Sequential Convex Programming for Aerocapture Trajectory Optimization
(Jens Ratajczak)



Training Neural Networks Without Gradients: a scalable ADMM approach
(K. Aditi)

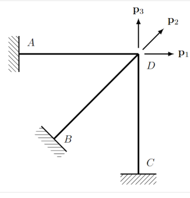
- In ADMM, Lagrange multiplier is added to the constraints.
- There exists a similar formulation known as Bregman iteration where Lagrange multiplier is added to the objective term.

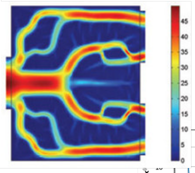
$$\min_{\{W_l\},\{z_l\},\{a_l\}} l(z_L, y) + \beta_L \|z_L - W_L a_{L-1}\|^2 + \langle z_L, \lambda \rangle + \sum_{l=1}^{L-1} \left(\gamma_l \|a_l - h_l(z_l)\|^2 + \beta_l \|z_l - W_l a_{l-1}\|^2 \right)$$

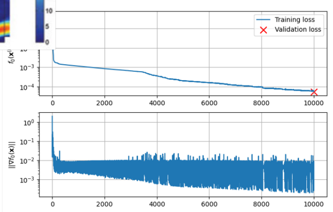
Student backgrounds:

- Applied Math (BS/MS, MS, PMD, PhD)
- Computer Science (MS, PhD)
- Electrical Engineering (MS, PhD)
- Aerospace Engineering (MS, PhD)

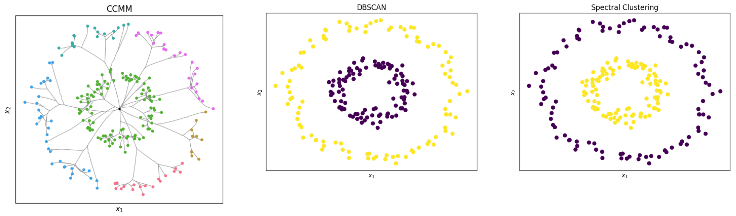
Method of Moving Asymptotes
(Grant Norman)







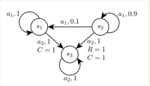
Convex Clustering though MM: An Efficient Algorithm to Perform Hierarchical Clustering
(Kal Parvanov, Bisman Singh)



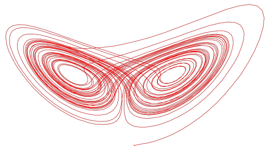
Column Generation Algorithms for Constrained POMDPs
(Tyler Becker and Qi Heng Ho)

Toy Problem

Solver	Value	Average Cost	Action Probabilities (initial belief)
CGCP	0.95	0.95	[0.05,0.95]
Optimal Deterministic	0.72	0.8	[1,0]



Convex Optimization for Fixed-Point Stability Analysis
(Morgan Byers)



Solving with PyDrake

```
prog = MathematicalProgram()

s = prog.NewIndeterminates(1, "s")[0]
c = prog.NewIndeterminates(1, "c")[0]
thetadot = prog.NewIndeterminates(1, "\dot\\theta")[0]

x = np.array([s, c, thetadot])

deg_V = 2
V = prog.NewFreePolynomial(Variables(x), deg_V).ToExpression()

constraint = prog.AddLinearConstraint(
    V.Substitute({s: 0, c: 1, thetadot: 0}) == 0
    )
```

mmWave Radar IM (low-rank + sparse)
(Yunxuan Wang)

$$\min_{\mathbf{I},X} \operatorname{rank}(\mathbf{I}) + \lambda \|\mathbf{X}_{\mathrm{RD}}\|_0 \quad \text{s.t.} \|\mathbf{Y} - \mathbf{I} - \mathbf{X}\|_F^2 < \delta,$$

$$\min_{\mathbf{I},X} \|\mathbf{I}\|_* + \lambda \|\mathbf{X}_{\mathrm{RD}}\|_1 \quad \text{s.t.} \|\mathbf{Y} - \mathbf{I} - \mathbf{X}\|_F^2 < \delta.$$

Deriving the Lovasz Extension for Submodular Minimization
(Rick Nueve)

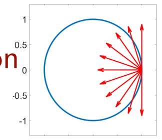
Theorem (1 : Greedy algorithm for submodular and base polyhedra, [1])

Let F be a submodular function such that $F(\emptyset) = 0$. Let $w \in \mathbb{R}^n$, with components ordered in decreasing order i.e., $w_{j_1} \geqslant \cdots \geqslant w_{j_n}$ and define $s_{j_k} = F(\{j_1, \ldots, j_k\}) - F(\{j_1, \ldots, j_{k-1}\})$. Then $s \in B(F)$ (i.e., a point from the base-polyhedra) and,

1. *if $w \in \mathbb{R}_+^n$, s is a maximizer of $\max_{s \in P(F)} w^T s$ (i.e., the support function over $P(F)$); moreover $\max_{s\in P(F)} w^T s = f(w)$.*
2. *s is a maximizer of $\max_{s\in B(F)} w^T s$, and $\max_{s\in B(F)} w^T s = f(w)$,*

Safe Feasibility Guided MPC for stochastic hybrid systems
(Tyler Becker and Qi Heng Ho)

Safe Gradient Flow for Convex Optimization
(Yiting Chen)



Proposition 1: $F_\alpha(x) \in T_{\mathcal{S}}(x)$, i.e. the gradient of the solution at the boundary points inside the $\mathcal{S}$ or matches with the tangent.