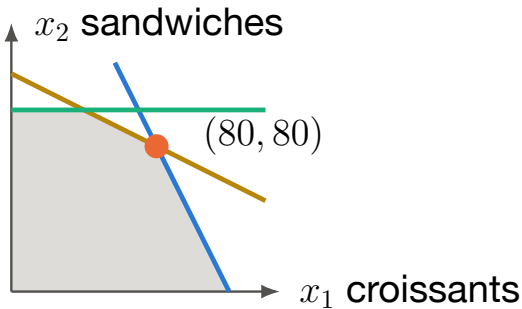
Second-order cone and semidefinite optimization

208424 Lecture 9

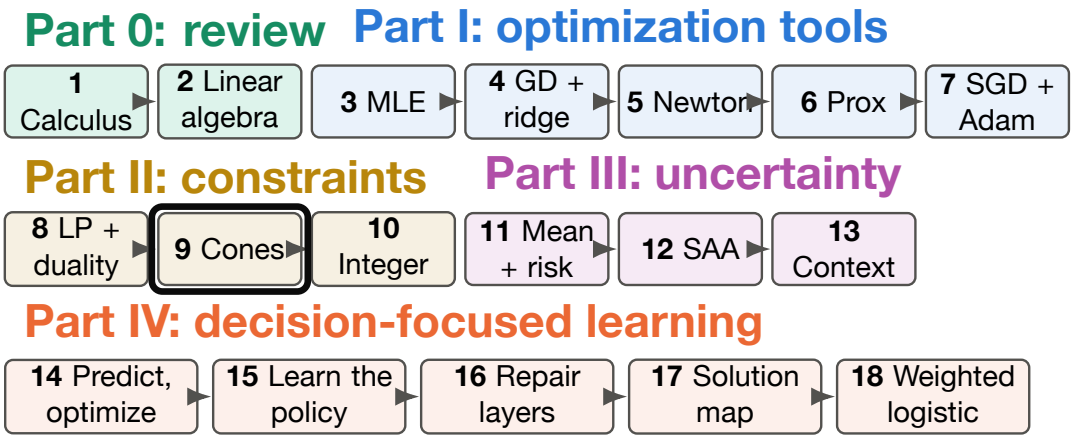
Lecture 8



In Lecture 8, every constraint was linear.

The best plan was at a vertex.

Course map


Part 0: review
Part I: optimization tools
1 Calculus
2 Linear algebra
3 MLE
4 GD + ridge
5 Newton
6 Prox
7 SGD + Adam
Part II: constraints
Part III: uncertainty
8 LP + duality
9 Cones
10 Integer
11 Mean + risk
12 SAA
13 Context
Part IV: decision-focused learning
14 Predict, optimize
15 Learn the policy
16 Repair layers
17 Solution map
18 Weighted logistic

Lecture 9: Second-order cone and semidefinite optimization

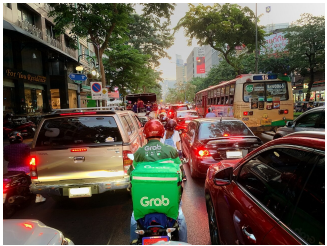
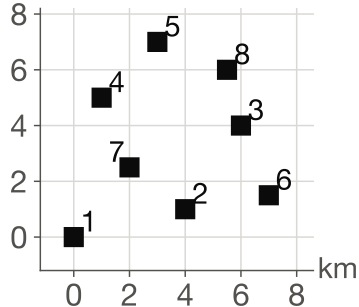


Photo: Chainwit., CC BY-SA 4.0, Wikimedia Commons

Where should a company put its delivery hub?

We add constraints on distances and matrices.

Eight cafés

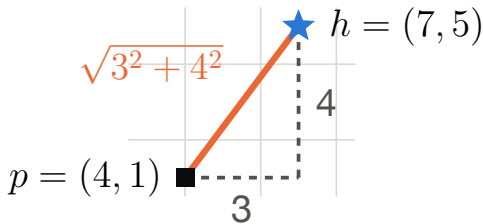


A company delivers to 8 cafés.

All deliveries start from one hub.

Where should the company put the hub?

Distance as a length



The distance from hub h to café p is $\|h - p\|$.

$$\|h - p\| = \sqrt{(h_1 - 4)^2 + (h_2 - 1)^2}$$

From $(7, 5)$ to $(4, 1)$, the distance is 5 km.

A distance constraint

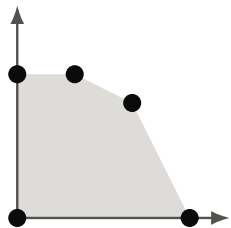


The hub must be at most 3 km from the café.

$$\|h - p\| \leq 3$$

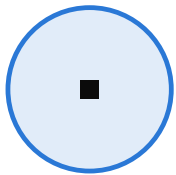
The feasible region is a disk.

A nonlinear constraint



linear constraints

$2x_1 + x_2 \leq 240$



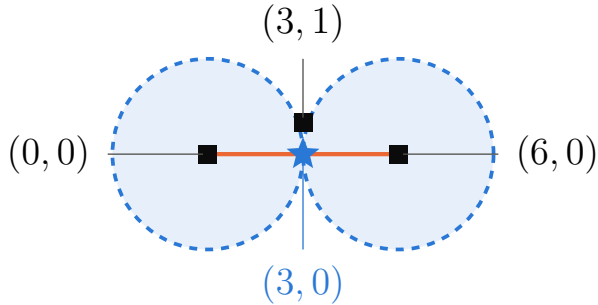
a distance constraint

$\sqrt{(h_1 - 4)^2 + (h_2 - 1)^2} \leq 3$

Because of the squares, $\|h - p\| \leq 3$ is nonlinear.

A disk has no straight edges and no vertices.

Three cafés

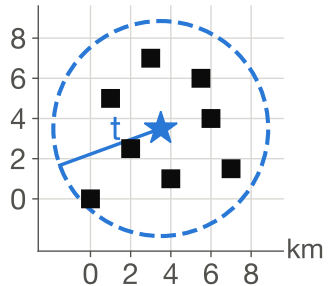


We want to minimize the maximum distance.

At $(3, 0)$, the maximum distance is 3 km.

Only $(3, 0)$ is within 3 km of both $(0, 0)$ and $(6, 0)$.

The hub problem

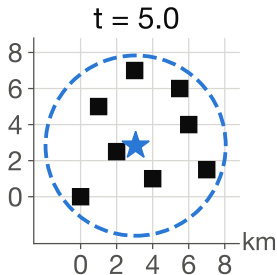


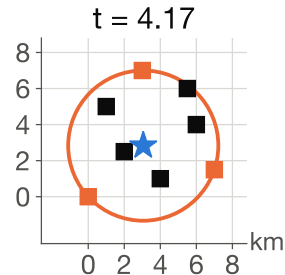
Café i is at the point p_i.

Let t be an upper bound on every distance.

minimize t subject to $\|h - p_i\| \leq t, \ \ i = 1, \ldots, 8$

The smallest t



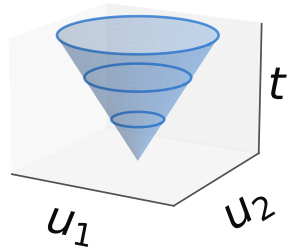


The decision variables are h_1, h_2 and t.

If t is above every distance, we can lower t.

So at the minimum, t is the maximum distance.

The second-order cone



The points (u_1, u_2, t) with $\sqrt{u_1^2 + u_2^2} \leq t$ form a cone.

It is called the second-order cone.

Café i's constraint means $(h - p_i,\ t)$ is in the cone.

Cone constraints

$$\underbrace{\|h - p_i\|}_{\text{the length of a linear expression}} \leq \underbrace{t}_{\text{a linear expression}}$$

This is a second-order cone constraint.

We also call it a cone constraint.

A second-order cone problem (SOCP)

An SOCP has a linear objective.

Its constraints are linear or cone constraints.

minimize t subject to $\|h - p_i\| \leq t, \;\; i = 1, \ldots, 8$

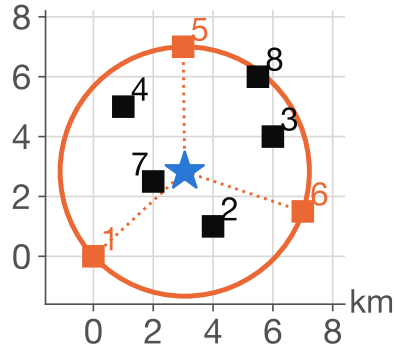
The hub in Python

```
import numpy as np, cvxpy as cp
cafes = np.array([[0, 0], [4, 1], [6, 4], [1, 5],
   [3, 7], [7, 1.5], [2, 2.5], [5.5, 6]])
h, t = cp.Variable(2), cp.Variable()
cons = [cp.norm(h - p) <= t for p in cafes]
cp.Problem(cp.Minimize(t), cons).solve()
```

cvxpy is a Python package for problems like this.

It passes the problem to a solver.

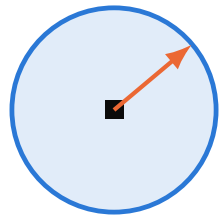
The best hub



The best hub is at $(3.05,\ 2.83)$.

The three farthest cafés are 4.17 km away.

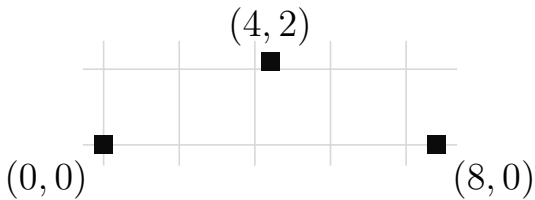
Summary: cone constraints



A distance constraint is a cone constraint.

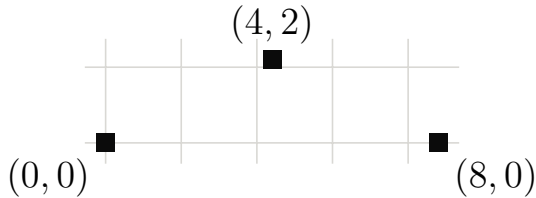
The hub problem is an SOCP.

Checkpoint 1



Where is the best hub for these three cafés?

Checkpoint 1

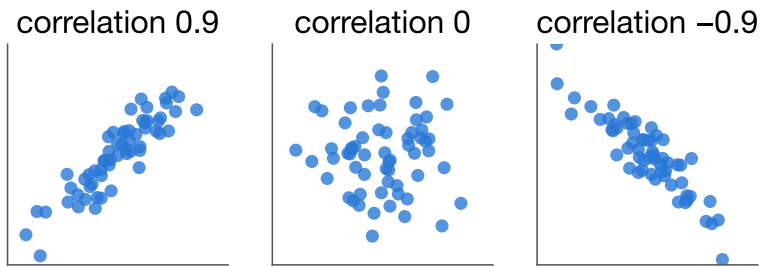


Where is the best hub for these three cafés?

The best hub is at $(4, 0)$.

Two cafés are 4 km away, and $(4, 2)$ is 2 km away.

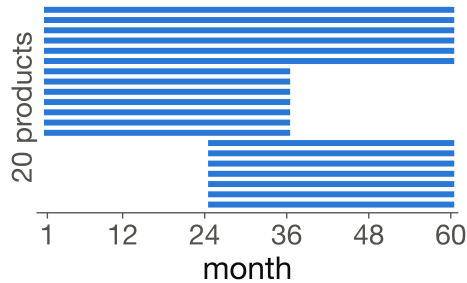
Correlation



A correlation is a number between -1 and 1.

It measures how two variables move together.

Sales of 20 products

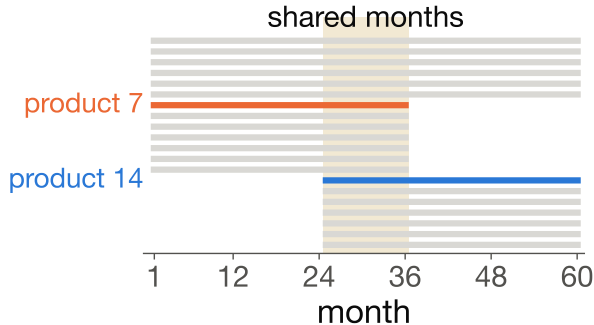


A shop sells 20 products.

Some products started late or stopped early.

The manager wants the correlation of each pair.

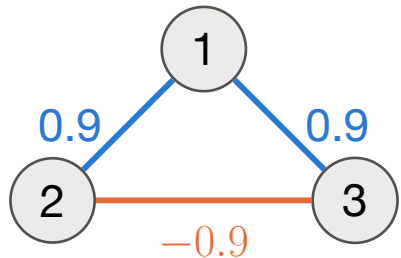
Each pair uses its shared months



Each correlation uses only the shared months.

Different pairs use different months.

Three products



Products 2 and 3 both move with product 1.

But products 2 and 3 move in opposite directions.

Can real data give these three correlations?

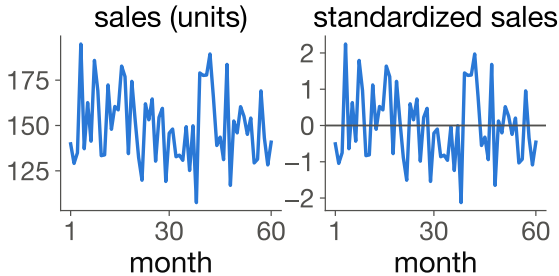
A correlation matrix

$$C = \begin{pmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & -0.9 \\ 0.9 & -0.9 & 1 \end{pmatrix}$$

C_{ij} is the correlation of products i and j.

Every diagonal entry is 1.

Standardized sales



We subtract the mean and divide by the SD.

Standardized sales have mean 0 and SD 1.

We call them z_1, z_2 and z_3.

The variance of a combination

Let C be the correlation matrix of z_1, z_2 and z_3.

The variance of $v_1 z_1 + v_2 z_2 + v_3 z_3$ is $v^\top C v$.

A variance is an average of squares, so $v^\top C v \geq 0$.

A negative variance

$v = (1, -1, -1)$

	1	-1	-1
1	1	-0.9	-0.9
-1	-0.9	1	-0.9
-1	-0.9	-0.9	1

$\longleftarrow v_1 v_3 C_{13} = 1 \times (-1) \times 0.9$

$v^\top C v$ adds the nine terms $v_i v_j C_{ij}$.

$v^\top C v = 3 - 6 \times 0.9 = -2.4$, which is negative.

Data from the same months can never give this C.

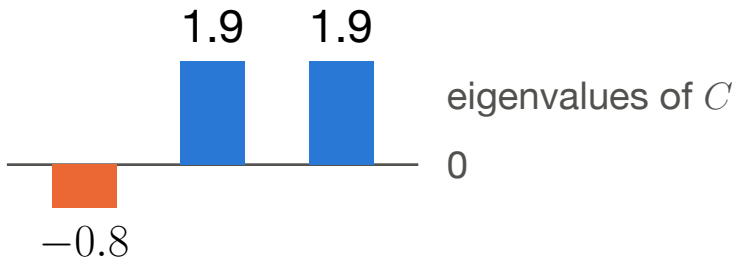
Positive semidefinite matrices

$C \succeq 0$ means $v^\top C v \geq 0$ for every v.

We say C is positive semidefinite (Lecture 2).

Our C is not positive semidefinite.

The eigenvalue test



For a symmetric C, $C \succeq 0$ exactly when every eigenvalue of C is ≥ 0.

One eigenvalue of our C is negative.

A valid correlation matrix

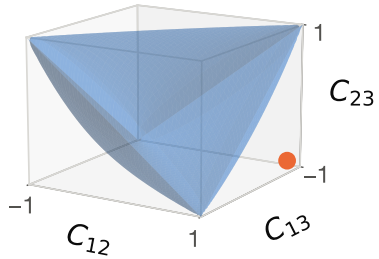
$$\begin{pmatrix} 1 & C_{12} & C_{13} \\ C_{12} & 1 & C_{23} \\ C_{13} & C_{23} & 1 \end{pmatrix}$$

A valid correlation matrix is symmetric.

Every diagonal entry is 1, and $C \succeq 0$.

Data from the same months give a valid matrix.

Valid matrices for three products

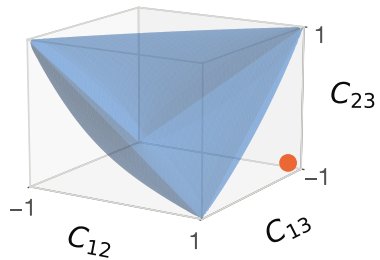


Each point is a choice of C_{12}, C_{13} and C_{23}.

The valid matrices form this shape.

The point $(0.9,\ 0.9,\ -0.9)$ lies outside the shape.

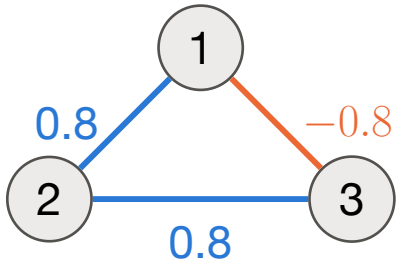
Summary: valid correlation matrices



A valid correlation matrix satisfies $C \succeq 0$.

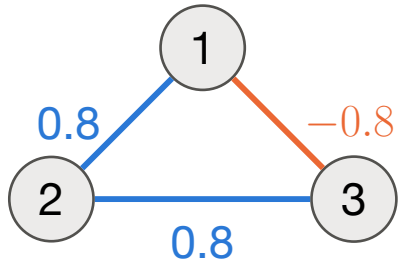
So every eigenvalue of C is ≥ 0.

Checkpoint 2



Is $v^\top Cv$ negative for $v = (1, -1, 1)$?

Checkpoint 2

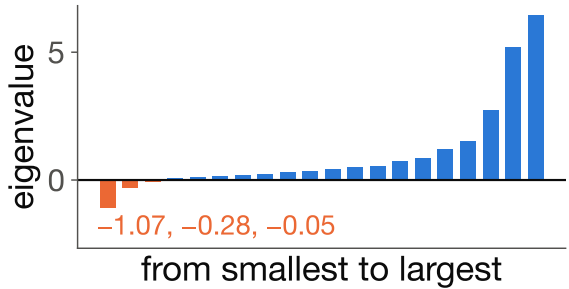


Is $v^\top Cv$ negative for $v = (1, -1, 1)$?

Yes. $v^\top Cv = 3 - 6 \times 0.8 = -1.8$.

So this correlation matrix is not valid.

The 20-product estimate



$\hat{C}$ is the matrix of the 190 estimated correlations.

Its correlations come from different months.

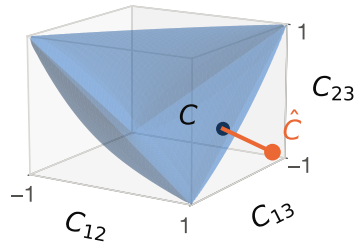
Three eigenvalues are negative, so $\hat{C}$ is not valid.

The distance between two matrices

$$\|C - \hat{C}\| = \sqrt{\sum_{i,j} (C_{ij} - \hat{C}_{ij})^2}$$

This is Lecture 2's length, using all entries.

The nearest valid correlation matrix



We want the valid C that is closest to $\hat{C}$.

minimize $\|C - \hat{C}\|$

subject to $C \succeq 0$ and every $C_{jj} = 1$

A semidefinite problem (SDP)

$x_1 + x_2 \leq 4$	$\|h - p\| \leq t$	$C \succeq 0$
linear	cone	semidefinite

$C \succeq 0$ is a semidefinite constraint.

An SDP has a linear objective.

Its constraints are linear, cone or semidefinite.

An upper bound on the distance

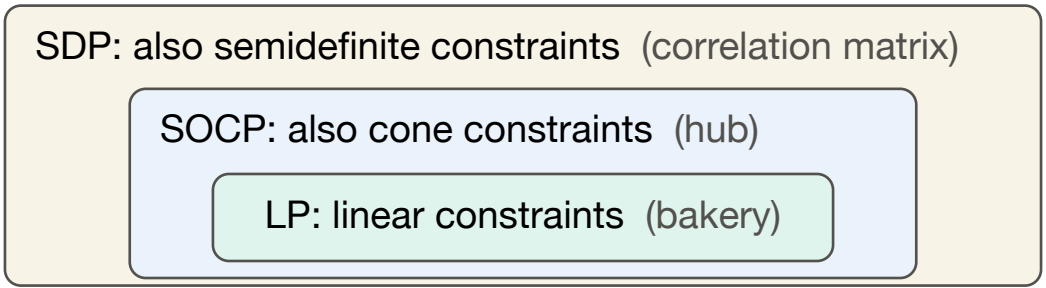
minimize s subject to $\|C - \hat{C}\| \leq s, \quad C \succeq 0, \quad C_{jj} = 1$

cone semidefinite linear

As with the hub, s is an upper bound on $\|C - \hat{C}\|$.

So this problem is an SDP.

Three kinds of problems



Every LP is an SOCP.

Every SOCP is an SDP.

Three products in Python

```
import numpy as np, cvxpy as cp
Chat = np.array([[1, .9, .9], [.9, 1, -.9],
   [.9, -.9, 1]])
C = cp.Variable((3, 3), symmetric=True)
cons = [C >> 0, cp.diag(C) == 1]
gap = cp.norm(C - Chat, "fro")
cp.Problem(cp.Minimize(gap), cons).solve()
np.linalg.eigvalsh(C.value)  # [0, 1.5, 1.5]
```

This code finds the nearest valid correlation matrix.

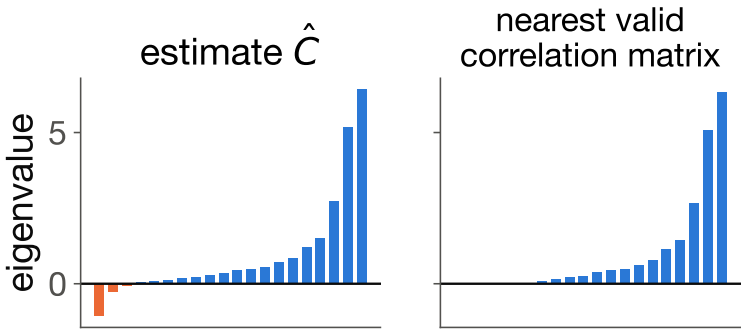
In cvxpy, `C >> 0` means $C \succeq 0$.

Three products: before and after

	estimate $\hat{C}$	nearest valid correlation matrix
C_{12}	0.9	0.5
C_{13}	0.9	0.5
C_{23}	-0.9	-0.5
eigenvalues	-0.8, 1.9, 1.9	0, 1.5, 1.5

Now every eigenvalue is ≥ 0, so C is valid.

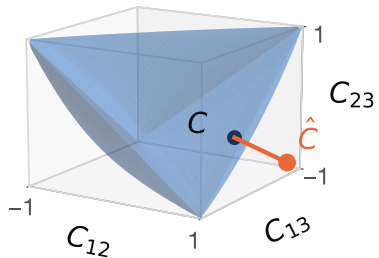
20 products: before and after



The smallest eigenvalue increases from -1.07 to 0.

The largest change in a correlation is 0.21.

Summary: semidefinite constraints



$C \succeq 0$ is a semidefinite constraint.

The nearest valid correlation matrix is an SDP.

Checkpoint 3

Which kind of constraint is $C_{jj} = 1$?

linear

cone

semidefinite

Checkpoint 3

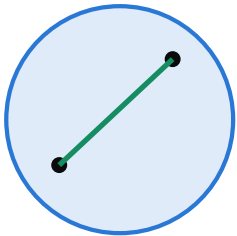
Which kind of constraint is $C_{jj} = 1$?

linear | cone | semidefinite

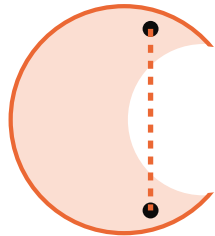
$C_{jj} = 1$ is a linear constraint.

Each diagonal entry of C must be 1.

Convex regions



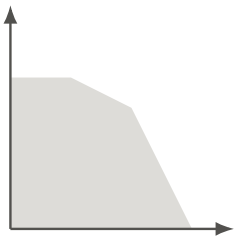
convex



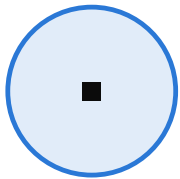
not convex

A region is convex if, for any two of its points, the line segment between them lies in the region.

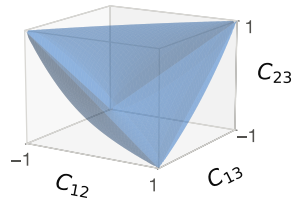
Our feasible regions are convex



linear



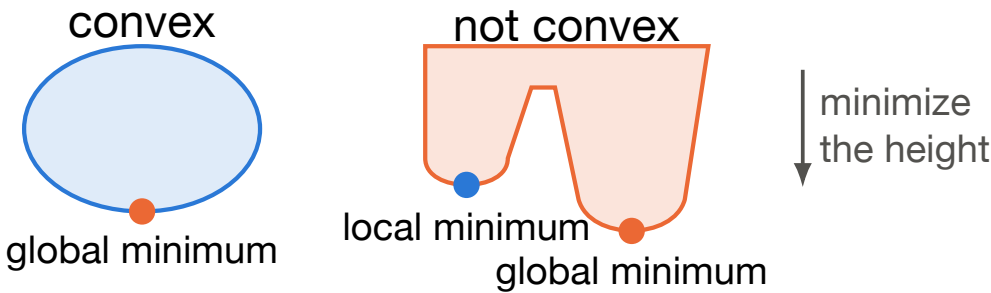
cone



semidefinite

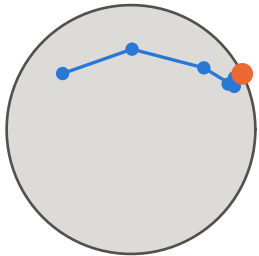
Linear, cone and semidefinite constraints
all give convex feasible regions.

A local minimum is the global minimum



With a linear objective and a convex region, every local minimum is the global minimum.

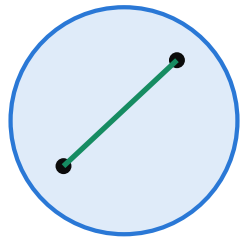
How a solver works



Interior-point methods solve SOCPs and SDPs.

Each step stays inside the feasible region.

Summary: convexity

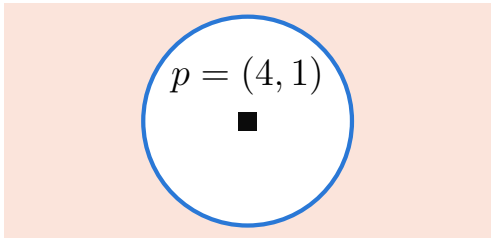


LPs, SOCPs and SDPs

have convex feasible regions.

So a local minimum is the global minimum.

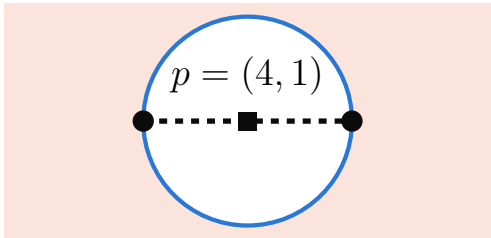
Checkpoint 4



$\|h - p\| \geq 3$

Is the region $\|h - p\| \geq 3$ convex?

Checkpoint 4



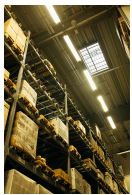
$\|h - p\| \geq 3$

Is the region $\|h - p\| \geq 3$ convex?

No. $(1, 1)$ and $(7, 1)$ are in the region.

The midpoint $(4, 1)$ of their line segment is not.

Where SOCPs and SDPs are used



warehouses



portfolios



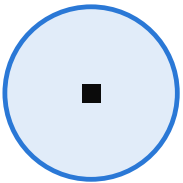
bridges

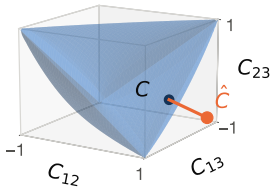


antennas

Photos (Wikimedia Commons): Shixart1985, CC BY 2.0; Francisco Gonzalez, CC BY 2.0; Daniel Schwen, CC BY-SA 3.0; Wikideas1, CC0.

Lecture summary

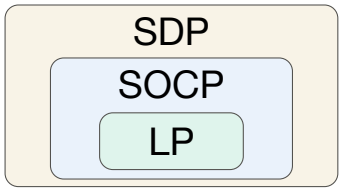


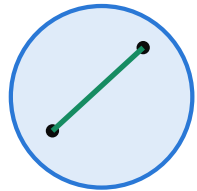


A distance constraint is a cone constraint.

The nearest valid correlation matrix is an SDP.

Lecture summary (continued)

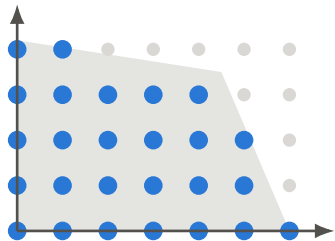




Every LP is an SOCP, and every SOCP is an SDP.

With a linear objective and a convex region,
a local minimum is the global minimum.

Next lecture



In Lecture 10, some decision variables are integers.

Exercises

1. Write “at most 5 km from $(2, 3)$” as a constraint.
2. Find the best hub for $(0, 0)$, $(4, 0)$ and $(2, 4)$.
3. Every correlation is 0.5. Check that $C \succeq 0$.