Semianfrite programming: What happens when you want to find
\[
\max_x f(x, y)?
\]
One can form \( g(x) = \min_y f(x, y) \), but this function might fail to be smooth:

Also, what if one has a continuous family of constraints:
\[
g(x, z) \leq 0 \quad \text{(before, } z \in \text{ finite set}).
\]

Typical examples:
- Distance between 2 sets \( d(X, Y) = \max_{x \in X} \min_{y \in Y} |x - y| \) + symmetry term.
- Find ellipsoid of smallest area containing some body \( B \) \( \det X \rightarrow \max \): \( X \succeq 0 \),
  \[
  (X_{\bar{z}, \bar{z}}) - I \leq 0 \quad \text{for } z \in B.
  \]

Recall
\[
\text{Lemma: } \text{If } K \subset \mathbb{R}^d \text{ is compact, and } f(h, z) < 0 \text{ for all } z \in B
\]
(was what we used to call \( \mathcal{K}^* \)) is empty, then \( O \in \text{conv}(K) \).

Thus: If in
\[
\min_{x \in X} \max_{y \in Y} f(x, y) \quad \text{subj to } \quad g(x, z) \leq 0, \ z \in Z
\]
(here \( x \in X \) is open, \( Y, Z \) are compact, \( f, g \), \( \nabla_x f(x, y) \) \( \nabla_x g(x, z) \) are continuous), and \( x \) is a local minimizer, then there is a finite number of \( y_j \), \( z_k \) at most \( n+1 \) altogether:
\[
f_x(z_j y_j) = \max_y f(x, y_j) \quad g(x, z) = 0
\]
for all \( j \) for all \( k \)
and
\[
\sum_i \lambda_i \nabla f(x_i y_i) + \sum_k \mu_k \nabla g(x_k y_k) = 0 \quad \text{for some } \lambda, \mu \geq 0
\]
not all \( = 0 \).
Proof: \[ \varphi(x) = \max_{y \in \mathbb{R}^n} f(x,y); \quad \psi(x) = \max_{z \in \mathbb{R}^m} g(x,z) \]
\[ \Theta(x) = \max \left( \varphi(x) - \psi(x), 0 \right) \]
If \( x^* \) is local minimizer (i.e., in a vicinity of \( x^* \), for any \( x: g(x,z) \leq 0 \) for all \( z \Rightarrow \varphi(x) > \psi(x) ) \) then \( \Theta(x) > 0 \)

Define active sets \( \mathcal{Y}_* = \{ y: f(x^*,y) = \max_{y \in \mathbb{R}^n} f(x,y) \} \)
\[ \mathcal{Z}_x = \{ z: g(x^*,z) = 0 \} \]

If for some \( h \in \mathbb{R}^n \) and any \( y \in \mathcal{Y}_*, z \in \mathcal{Z}_x \)
\[ (\nabla_y f(x^*,y), h) < 0 \quad \text{and} \quad (\nabla_z g(x^*,z), h) < 0 \]
then \( x_t^* + \mathbf{t} \mathbf{h} \) will have
\[ f(x_t^*+\mathbf{t}\mathbf{h},y) < f(y^*) \]
for small \( t \) and all \( y \) (this is true over some \( [0, t(y)] \) for all \( y \), whether in \( Y_* \), or \( Y-Y_* \)) hence true for some close \( y \)'s; by compactness can choose finite over, and smallest of \( t(y) > 0 \), and some way \( g(x^*_t+\mathbf{t}\mathbf{h},z) \leq 0 \) for small \( t, \) all \( z \).
This contradicts assumption on \( x^* \)
\[ \bigcup \left\{ h: (\nabla_y f(x^*,y), h) < 0 \right\} \]
\[ \bigcup_{y \in \mathcal{Y}_*} \bigcup_{z \in \mathcal{Z}_x} \bigcup \left\{ h: (\nabla_z g(x^*,z), h) < 0 \right\} = \emptyset. \]

So: \( 0 \in \text{conv. hull of gradients} \quad \nabla_y f(x^*,y), Y \in \mathcal{Y}_*; \quad \nabla_z g(x^*,z), Z \in \mathcal{Z}_x \)

But, by Radon's theorem, one can find \( \leq n+1 \) vectors among those, whose convex hull is \( 0 \Rightarrow \) conclusion
Example. Thug’s Inequality. 

Let $S \subseteq \mathbb{R}^n$ be a compact set. 

Define $D(S) = \sup_{x,y \in S} |x - y|$. 

$\mathbf{R}(S) =$ smallest radius of ball containing $S$. 

i.e. $\mathbf{R}(S) = \min \max_{x,y \in S} \frac{1}{2} |x - y|^2$ 

Then $D(S) \geq \sqrt{\frac{2(n+1)}{n}} \mathbf{R}(S)$. 

Take $\varphi(x) = \max_{y \in S} |x - y|^2$. $\varphi$ is max of convex function coercive $\implies$ unique minimizer $x_\star$. 

By the above, there are $n+1$ points in $S$: $|x - y|^2 = \mathbf{R}^2(S)$ and for some $\lambda \geq 0$ 

$$\sum_{j} \lambda_j \mathbf{y}_j \cdot (x - y_j) = \sum_{j} \lambda_j (x - y_j)^2 = 2 \sum_{j} \lambda_j (x - y_j) = 0$$

Can rescale and shift, so that $x_\star = 0, \sum \lambda_j = 1, \lambda \geq 0, \sum \lambda_j y_j = 0$ 

Then 

$$\sum_{j,k} \lambda_j \lambda_k \|y_j - y_k\|^2 = \sum_{j} \lambda_j \|y_j\|^2 + \sum_{j,k} \lambda_j \lambda_k \|y_j - y_k\|^2 - 2 \sum_{j,k} \lambda_j \lambda_k (y_j, y_k)$$

$$= 2 \sum_{j} \lambda_j \|y_j\|^2 - 2 \sum_{j} \lambda_j (y_j, y_k) = 2 \mathbf{R}^2(S)$$

$$\implies 2 \mathbf{R}^2(S) = \sum_{j,k} \lambda_j \lambda_k \|y_j - y_k\|^2 \leq \sum_{j,k} \lambda_j \lambda_k \|y_j - y_k\|^2$$

$$\sum_{j,k} \lambda_j \lambda_k = 1 \implies \sum_{j,k} \lambda_j \lambda_k \|y_j - y_k\|^2 \leq \sum_{j,k} \lambda_j \lambda_k \cdot D^2(S) \implies \sum_{j,k} \lambda_j \lambda_k \cdot D^2(S) \leq 1 - \frac{1}{k} \text{ by Cauchy-Schwarz.}$$

$$(\sum_{j,k} \lambda_j^2 \cdot 1^2 \leq (\sum \lambda_j^2)^2 = 1 \implies \sum \lambda_j^2 \geq 1 - \frac{1}{k})$$.

So, $\mathbf{R}^2(S) \leq \left(\frac{k-1}{k}\right)^2 D(S)$

Inequality exact: see regular simplex ...
Last example: Best approximation for $x^n$ on $[-1,1]$ in $L_\infty$-norm, among polynomials of degree $< n$.

In other words, we solve

$$\begin{align*}
\mathbf{z} & \rightarrow \min \quad \text{subject to} \\
-\mathbf{z} \leq x^m + \sum_{k=1}^n a_k x^k \leq \mathbf{z} \quad \text{for } x \in [-1,1].
\end{align*}$$

Theorem implies there are $\leq n+1$ points, where the constraints are active, for optimal $\mathbf{a}^*_n$ (we call $P^*$ corresponding polynomial).

Define these points $t_k, t'_k \in [-1,1]$ where $P^*_n(t_k) = 2$; $P^*_n(t'_k) = -2$.

1. If $t_k = t'_k \in (-1,1)$, then multiplicity of $(P^*_n + \mathbf{z})$ at that point is even. Call these multiplicities $m_k, m'_k$.

2. Sum of $m_k$ is $n$; otherwise, $P^* - \sum \left\{ m_k (t_k - t) \right\} \mathbf{e}_2$ fits into a smaller $\mathbf{z}$-box. Can use $\min\{m_k, 2\} \rightarrow 1$ proves $m_k = 2$.

3. One has $-1$ and $+1$ among $t_k, t'_k$. (Otherwise, taking $P^*_n(t - a)$ can be used to show that $\Sigma m_k, n \Sigma m'_k$ is less than $n$, contradiction.)

OK: All that implies that $\mathbf{P} = \frac{1}{2} P^*_n$ is such that the leading coeff. is $x^{n/2}$; and $\mathbf{P} + 1$ has $n$ root $\in [-1,1]$. Take $(\mathbf{P}'_n)^2 (x^2 - 1)$ — it has same roots as $(\mathbf{P}^*_n)^2 - 1$, and lead. coefficients are $n^2/2^2$ vs $1/2^2$. So
\[ n \left( \tilde{p}^2 - 1 \right) = (\tilde{p}')^2 (x^2 - 1) \quad \text{i.e. } \tilde{p} \text{ solves} \]

\[ \frac{p'}{\sqrt{1-p^2}} = \frac{n}{\sqrt{1-x^2}} \quad \Rightarrow \quad \int \frac{dp}{\sqrt{1-p^2}} = n \int \frac{dx}{\sqrt{1-x^2}} \]

\[ \cos \tilde{p} = n \cos x + c \ldots \]