**Convex Polytope**

A polytope is defined as the convex hull of a finite set of points. In dimension $d$, the equation $\sum a_i x = b$ defines a $d-1$-dimensional object.

- Bounded polyhedron
- Intersection of finite # of half-spaces: $\sum a_i x \leq b_i$

A is $m \times d$, and $x \geq 0$. For example, $\sum a_i x = b$ implies $x_i = \frac{b_i}{a_i}$. The intersection of $r$ hyperplanes gives you a $d-r$-dimensional object.

A polytope is $d$-dimensional, and faces can be $d-1$ or $d-2$ dimensional.
Example: unit (hyper)-cube

(solid translation version of $B_2 = \text{ball} \ | \ |x| = 1 \}$)

- $B_n$-ball $\{x : \|x\|_2 = 1\}$
- $\text{conv}(\sum_{i=1}^{d} e_i i)$
- simplex $\Delta^d$: $d$ vertices $\rightarrow e_i^*$
  $d$ facets $\rightarrow x_i = 0$

- in structured sum, we have $\bigoplus_{i=1}^{d} x_i$ (product structure vs "cube $\Delta^d$"")

In simplex: $\|x\|_2 = 1$ always but $\|x\|_1$ as small as $\frac{1}{\sqrt{d}}$

$\|\frac{1}{d} 1\|_1 = 1$, $\|\frac{1}{d} 1\|_2 = \sqrt{\frac{d^2 + \cdots + 1}{d}} = \frac{1}{\sqrt{d}}$

* pitfalls of low-dimensional intuition vs high dimension

$\text{dist from origin} = \frac{\|d\|_1}{\sqrt{d}} \approx \sqrt{\frac{d}{2}}$
Affine invariance of FW:

recall (lecture 0) \( \hat{M} \xrightarrow{\sim} \hat{M} \times \hat{M} \)

\( \hat{S}(x) = \frac{\hat{S}(\hat{A}x)}{\hat{a} \cdot \hat{M}} \)

running FW on \( \hat{M} \hat{S} \hat{F} \)

equivalent to running on \( M \hat{S} \hat{F} \)

\( x_6 = \hat{A} \hat{a} \hat{x}_6 \)

well let \( \hat{M}_i = \Delta i \hat{F} \) let \( \hat{A} \) be marginalization operation

\( M_0(y_c) = \sum_y \alpha(y_c) = \alpha^T \mathbf{y} \mathbf{y}_c \)

Marginal polytope:

* need to decide what "parameterization" of marginals you want to use...

  e.g. 1) simplest: node's edge marginals at possible values in (redundant)
Getting a representation of marginal polytope:

1) \( \mu_i(y_i) \rightarrow 1 \cdot k \) variables
2) Minimal representation \( \rightarrow |V| \cdot (k-1) + |E| \cdot (k-1)^2 \)
3) Use \( \mu_i(y_i) \) with maximal cliques

Fournier-Metzelen elimination:

Start with \( \Delta_{\lambda y} \rightarrow \mathcal{E} (y) = a \) and \( a(y) = 0 \) \( \forall y \in \mathcal{X} \)

If \( q^{[x]} = b \) \( \quad q^{[y, x]} > b \)

if \( q^{[x]} = b \) \( \quad q^{[y, x]} < b \)

Add constraint: \( \Delta = a \)

to get \( M \), the marginal polytope

You want to eliminate \( \alpha \) variable

from the description

First, make \( \Delta = a \) as inequalities \( \Delta = a + \Delta \alpha \leq a \)

Fournier-Metzelen elimination: if \( q^{[x]} < 0 \)
\[
\frac{\sum_{i} a_i x_i}{\sum_{i} b_i} = x \quad \frac{\sum_{i} b_i - \sum_{i} a_i x_i}{\sum_{i} b_i - \sum_{i} a_i x_i} \quad \text{for } \sum_{i} b_i \geq \sum_{i} a_i x_i
\]

**Inequalities**

\[
\sum_{i} a_i x_i \leq \sum_{i} b_i \quad \Rightarrow \quad \text{No new inequalities}
\]

**Quadratic explosion**

\[
\sum_{i} b_i - \sum_{i} a_i x_i < 0
\]

**Remainder**

\[
f(x) = f(a) + (x-a)^T \nabla f(a) + \frac{1}{2} (x-a)^T H(f(a))(x-a)
\]

**Taylor**

\[
x = x + \gamma (s-x)
\]

**CG**

\[
CG = \sup_{s \in \mathcal{S}} (s-x)^T H(s-x)
\]

**For SVMstruct**

Hessian \( H = \sum_{i=1}^{\Delta x_i} \Delta x_i A \)

\[
\Rightarrow CG \leq \sup_{s \in \mathcal{S}} (s-x)^T H(s-x)
\]

**Teaching Page 5**
Let's compute $C_5 \leq \frac{4R^2}{\lambda}$ and let's use $1/\lambda$. Let $\text{diam}(M) = 2R$.

Lipschitz constant in $\ell_2$-norm = largest eigenvalue of Hessian

$$\lambda A A = \sum_{i,j} \langle \nabla_i f(y), \nabla_j f(y) \rangle_{\lambda A}$$

say e.g., $\langle \nabla_i f(y), \nabla_j f(y) \rangle \approx 1$ for lots of outputs

$$\lambda A A = 1 \cdot \lambda$$

Here, our huge bound for $C_5$ becomes useless.

Exercise to the reader: If you use $\|\cdot\|_{\ell_2}$ block-norm on $M$

i.e., $\|x\|_{\ell_2} = \max_{i=1,n} |x_i|$ for $x = [x_1, \ldots, x_n]$, then

$\text{diam}_{\|\cdot\|_{\ell_2}}(M)^2 = 2$

Let $x_1, \ldots, x_n$ be such that

$\|x\|_{\ell_2} = \frac{2R}{\lambda}$
Pointers:
- to learn more about polytopes:
- to learn more about the marginal polytope:
  - see Section 3.4 of:
    http://www.eecs.berkeley.edu/~wainwrig/Papers/WaiJor08_FTML.pdf

-> also, see p.80 Figure 4.1 for an example of fractional corners for the local consistency polytope (vs. the marginal polytope which only has integer vertices)