Originally, it was designed to handle mixed linear programming. Amazing!

0.1 Mixed-integer Zero-one Programming

0.1.1 Relaxation and Linearization

Consider a linear mixed-integer zero-one programming problem whose feasible region is given as follows:

$$X = \{(x,y) : \sum_{j=1}^{n} \alpha_{rj}x_j + \sum_{k=1}^{m} \gamma_{rk}y_k \geq \beta_r, \forall r \in [R],$$
$$x_j \in \{0,1\}, \forall j \in [n],$$
$$0 \leq y_k \leq 1, \forall k \in [m]\} \tag{1}$$

One classical approach involves relaxing the integrality constraint by simply treating the variables as continuous real values rather than integers, thereby allowing for an optimal but potentially infeasible solution. The Sherali-Adams reformulation-linearization is another approach.

For any $d \in [N]$, let us define the polynomial factors of degree d as

$$F_d(J_1, J_2) = (\prod_{j \in J_1} x_j)(\prod_{j \in J_2} (1 - x_j))$$

for each $J_1, J_2 \subseteq [N]$ such that $J_1 \cap J_2 = \phi$, and $|J_1 \cup J_2| = d$. Sepcially, $F_0(\phi, \phi) = 1$.

Sherali and Adams use each polynomial factor of differing degree to construct a hierarchy of relaxations. In the description of relaxation algorithm, we just fix a d. We multiply $F_d(J_1, J_2)$ on both sides and remove the integerality constraints, but, to some extent, hidenly add constraints $x_i(x_i - 1) = 0$. Then the programming problem (1) would be updated to

$$(\sum_{j \in J_1} \alpha_{rj})F_d(J_1, J_2) - \beta_r F_d(J_1, J_2) + \sum_{j \in [N] \backslash (J_1 \cup J_2)} \alpha_{rj}F_{d+1}(J_1 + j, J_2) + \sum_{k=1}^{m} \gamma_{rk}y_k F_d(J_1, J_2) \geq 0, \forall r \in [R], \forall F_d. \tag{2}$$

Because we remove integerality constraints, it's suffice to restrict $0 \leq x, y \leq 1$. Thus, we have another two constraints

$$F_D(J_1, J_2) \geq 0, D = min(d+1, N), \forall F_D$$
$$F_d(J_1, J_2) \geq y_k F_d(J_1, J_2) \geq 0, \forall F_d, \forall k \in [m]. \tag{3}$$

Viewing the constraints above in expanded form as a sum of monomials, linearize them by substituting the following variables for the corresponding nonlinear terms $\forall J \subseteq [N], \forall k \in [m], w_J = \prod_{j \in J} x_j, v_{Jk} = y_k \prod_{j \in J} x_j$. Specially, $w_\phi = 1$. Furthermore, denoting by $f_d(J_1, J_2)$ and $f_d^k(J_1, J_2)$ the respective linearized forms of the polynomial expressions $F_d(J_1, J_2)$ and $y_k F_d(J_1, J_2)$.

The original problem of finding the optimal (X, Y) is transformed into a new problem of finding the optimal (W, V). Actually, W includes $w_j = x_j$ and V includes $v_{\phi k} = y_k$. For simplicity, we can denote the solution as (X, Y, W, V). We then need to demonstrate the validity of this relaxation.

0.1.2 Validity

Theorem 0.1. $Conv(X) \subseteq X_{P_n} \subseteq X_{P_{n-1}} \subseteq ... \subseteq X_{P_1} \subseteq X_{P_0} = X_0$ where X_{P_d} is the corresponding $(\mathbf{x}, \mathbf{y})$ to $X_d = \{(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})\}$. In particular, $X_{P_d} \cap \{(x, y) : x \text{ is binary}\} = X$ for all $d \in [N]$.

Proof. First, we show $X_{P_d} \subseteq X_{P_{d-1}}$.

$$F_d(J_1 + j, J_2) \geq 0, F_d(J_1, J_2 + j) \geq 0 \Rightarrow F_{d-1}(J_1, J_2) \geq 0$$

Then constraints (3) hold. For constraints (2), all terms are analogous except $\sum_{j \in N \setminus J_1 \cup J_2} \alpha_{r_j} F_{d+1}(J_1 + j, J_2)$. We choose pairs $(J_1 + t, J_2)$ and $(J_1, J_2 + t)$ where $t \notin J_1 \cup J_2$. Then $\sum_{j \in N \setminus J_1 \cup J_2 \cup t}(F_{d+2}(J_1 + t + j, J_2)) + \sum_{j \in N \setminus J_1 \cup J_2 \cup t}(F_{d+2}(J_1 + j, J_2 + t)) + F_{d+1}(J_1 + t, J_2) = \sum_{j \in N \setminus J_1 \cup J_2} \alpha_{r_j} F_{d+1}(J_1 + j, J_2)$. Then we can conclude $X_{P_d} \subseteq X_{P_{d-1}}$.

Second, let us show that $Conv(X) \subseteq X_{P_n}$. $X \subseteq X_{P_n}$ is trivial. And X_{P_n} is convex. Thus $Conv(X) \subseteq X_{P_n}$. In particular, $X = Conv(X) \cap \{(x, y) : x \ is \ binary\} = X_0 \cap \{(x, y) : x \ is \ binary\}$ because X_0 just removes integrality restrictions. Finally, it's suffice that $X_{P_d} \cap \{(x, y) : x \ is \ binary\} = X$ for all $d \in [N]$. $\square$

0.1.3 Convexity

Theorem 0.2. $X_{P_n} = Conv(X)$

0.2 Application on Proof System