A Restricted Dual Peaceman-Rachford Splitting Method for a Strengthened DNN Relaxation for QAP

Henry Wolkowicz
Dept. Comb. and Opt., University of Waterloo, Canada

Thur. May 19, 13:45-14:35 (BST) (08:45-09:35 EDT) 2022

At: Modern Techniques of Very Large Scale Optimization

joint with: Naomi Graham, Hao Hu, Jiyoung Im, Xinxin Li

1 Jilin Univ., China
Splitting methods: numerically hard, large scale problems (particularly successful for relaxations of hard nonlinear discrete optimization problems.)

We consider a Restricted Dual Peaceman-Rachford Splitting Method with strengthened bounds for a DNN relaxation; and we solve many NP-hard problems to (provable) optimality

Here: quadratic assignment problem, QAP, a fundamental HARD combinatorial optimization problems; QAP models many real-life problems such as facility location, VLSI design.
### Exploiting Structure/Novel

- We use facial reduction, FR, to obtain a natural splitting of variables into cone/polyhedral constraints.
- We modify the subproblems by adding redundant constraints.
- We use provable lower and upper bounds.
- We modify dual variable update by exploiting scaling.
- We present extensive numerical experiments. In many instances the DNN relaxation resulted in the global optimal solution of the QAP.

### Our Main Reference

A Restricted Dual Peaceman-Rachford Splitting Method for a Strengthened DNN Relaxation for QAP,

Graham/Hu/Im/Li/Wolkowicz in INFORMS J. Comput. 2022.

Further historical and current references are in the paper.
Facility Location

Given: \( n \) facilities and \( n \) locations; distance \( B_{st} \) between locations \( s, t \); flow \( A_{i,j} \) between facilities \( i, j \); location (building) cost \( C_{is} \) facility \( i \) in location \( s \).

\( X = X_{ij} \in \Pi \) permutation matrix unknown 0, 1 variables

\( X(\cdot) \in \mathbb{R}^{n^2}, \ n = 30 \) instances still considered hard

trace formulation, \( \langle Y, X \rangle = \text{trace}(YX^T) \)

minimize total flow and location costs

\[ p^*_\text{QAP} := \min_{X \in \Pi} \langle AXB - 2C, X \rangle, \]
\( X \in \mathbb{R}^{n \times n} ; \quad x = \text{vec}(X) \in \mathbb{R}^{n^2} \) (columnwise)

\[
Y := \begin{pmatrix}
1 \\
x
\end{pmatrix}
\begin{pmatrix}
1 & x^T
\end{pmatrix} \in \mathbb{S}^{n^2+1}
\]

**Block Representation**

Indexing the rows and columns of \( Y \) from 0 to \( n^2 \),

\[
Y = \begin{bmatrix}
Y_{00} & \bar{y}^T \\
\bar{y} & \bar{Y}
\end{bmatrix} , \quad \bar{y} = \begin{bmatrix}
Y_{(10)} \\
Y_{(20)} \\
\vdots \\
Y_{(n0)}
\end{bmatrix} , \quad \bar{Y} = xx^T = \begin{bmatrix}
\bar{Y}_{(11)} & \bar{Y}_{(12)} & \ldots & \bar{Y}_{(1n)} \\
\bar{Y}_{(21)} & \bar{Y}_{(22)} & \ldots & \bar{Y}_{(2n)} \\
\vdots & \vdots & \ddots & \vdots \\
\bar{Y}_{(n1)} & \bar{Y}_{(n2)} & \ldots & \bar{Y}_{(nn)}
\end{bmatrix}
\]

where

\[
\bar{Y}_{(ij)} = X_{i}X_{j}^T \in \mathbb{R}^{n \times n} , \forall i, j = 1, \ldots, n , \quad Y_{(j0)} \in \mathbb{R}^{n} , \forall j = 1, \ldots, n
\]
Reformulation of QAP

Lifted Objective

\[ L_Q := \begin{bmatrix} 0 & -(\text{vec}(C)^T) \\ -\text{vec}(C) & B \otimes A \end{bmatrix}, \quad (\otimes \text{ is Kronecker product}) \]

Lifted QAP

\[ p^*_{\text{QAP}} = \min \quad \langle AXB - 2C, X \rangle = \langle L_Q, Y \rangle \]

s.t. \[ Y := \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T \in S^{n^2+1}_+ \]

\[ X = \text{Mat}(x) \in \Pi, \]

where \( \text{Mat} = \text{vec}^* \), the adjoint transformation.
Characterizing Permutation Matrices

\[ \| Xe - e \|^2 = \| X^T e - e \|^2 = 0, \quad X \circ X = X, \quad X^T X = XX^T = I, \]

\( \circ \) is Hadamard product; \( e \) is vector of all ones.

Facial reduction FR using \( Xe - e = lXe - e = 0 \)

Let \( x = \text{vec}(X) \); \( y = \begin{pmatrix} 1 \\ x \end{pmatrix} \), \( Y = yy^T \)

We have \( lXe = (e^T \otimes I) \text{vec}(X) \) and

\[ Xe - e = 0 \quad \iff \quad y^T \begin{bmatrix} -e \\ (e^T \otimes I) \end{bmatrix} = 0 \]

\[ \iff \quad Y \begin{pmatrix} -e \\ (e^T \otimes I) \end{pmatrix} \begin{pmatrix} -e \\ (e^T \otimes I) \end{pmatrix}^T = 0 \]

We have an exposing matrix and can do FR, \( Y = \hat{V}R\hat{V}^T \).
After FR, $Y = \hat{V} R \hat{V}^T$; Primal-Dual Strong Duality Holds

smaller, greatly simplified, many constraints are redundant:

\[
\begin{align*}
\text{(SDP)} & \quad \min_{R} \quad \langle \hat{V}^T L_{Q} \hat{V}, R \rangle \\
\text{s.t.} & \quad G_{\bar{J}}(\hat{V} R \hat{V}^T) = u_0 \quad \text{(0-unit vector)} \\
& \quad R \in S_{+}^{(n-1)^2+1}.
\end{align*}
\]

$G_{\bar{J}}(\cdot)$ so-called \textit{gangster operator} ( [6] ZKRW’94)

fixes the elements in set $\bar{J}$. 

Details: $\hat{V}$, Facial Reduction

$\hat{Y}$

barycenter of set of feasible lifted $Y$ of rank one for the SDP relaxation;

$\hat{V} \in \mathbb{R}^{(n^2+1) \times ((n-1)^2+1)}$

have orthonormal columns that span the range of $\hat{Y}$ (explicit representation is available)

Minimal Face

every feasible $Y$ of the SDP relaxation is contained in the minimal face, $\mathcal{F}$ of $S_+^{n^2+1}$:

$$\mathcal{F} = \hat{V} S_+^{(n-1)^2+1} \hat{V}^T \Delta S_+^{n^2+1};$$

$$Y \in \mathcal{F}(\in \text{ri}(\mathcal{F})) \implies \text{Range}(Y) \subseteq (=) \text{Range}(\hat{V}),$$
linear map $G_J : \mathbb{S}^{n^2+1} \to \mathbb{R}^{|\bar{J}|}$; (shoots holes in the matrix)

By abuse of notation, also from $\mathbb{S}^{n^2+1}$ to $\mathbb{S}^{n^2+1}$, depending on the context:

$G_J : \mathbb{S}^{n^2+1} \to \mathbb{S}^{n^2+1}$, \quad \left[ G_J(Y) \right]_{ij} = \begin{cases} Y_{ij} & \text{if } (i, j) \in \bar{J} \text{ or } (j, i) \in \bar{J}, \\ 0 & \text{otherwise.} \end{cases}$

Gangster index set $\bar{J}$

union of the top left index (00) with set of indices $J$, $i < j$ in submatrix $\bar{Y} \in \mathbb{S}^{n^2}$:

(a) the off-diagonal elements in the $n$ diagonal blocks in $\bar{Y}$

(b) the diagonal elements in the off-diagonal blocks in $\bar{Y}$

Many of these are redundant; still used in subproblems.

The notation can be more tightly related to the $\{0, 1\}^{|\bar{J}|}$ with $1$ in the (00) entry.
Motivated by **natural splitting of variables** from FR

\( Y \) - in polyhedral constraints subproblem  
\( R \) - in SDP constraints subproblem

\[
\begin{align*}
\min_{R,Y} & \quad \langle L_Q, Y \rangle \\
\text{s.t.} & \quad Y = \hat{V} R \hat{V}^T \quad (\text{splitting}) \\
& \quad \mathcal{G}_J(Y) = u_0 \quad (\text{polyhedral}) \\
& \quad 0 \leq Y \leq 1 \quad (\text{polyhedral}) \\
& \quad R \succeq 0 \quad (\text{convex cone})
\end{align*}
\]

And, add redundant constraints to polyhedral and cone constraints.
Let $Y \in \mathbb{S}^{n^2+1}$ be blocked as above.

**Linearizations of Orthogonality: $XX^T = X^TX = I$**

- $b^\circ \text{diag}(Y) : \mathbb{S}^{n^2+1} \rightarrow \mathbb{S}^n$, sum of $n \times n$ diagonal blocks of $Y$:

  $$b^\circ \text{diag}(Y) := \sum_{k=1}^{n} Y_{(kk)} = I_n$$

- $o^\circ \text{diag}(Y) : \mathbb{S}^{n^2+1} \rightarrow \mathbb{S}^n$: traces of blocks $Y_{(ij)}$:

  $$o^\circ \text{diag}(Y) := \left( \text{trace} \left( Y_{(ij)} \right) \right) = I_n$$
More Redundant Constraints

Linearizations of $0, 1, X_{ij}^2 - X_{ij} = 0$

arrow($Y$) : $\mathbb{S}^{n^2+1} \rightarrow \mathbb{R}^{n^2+1}$; difference of first column and diagonal of $Y$:

$$\text{arrow}(Y) := (Y(:,1) - \text{diag}(Y)) = 0$$

trace constraint

By commutativity of the trace operator and $\hat{V}^T\hat{V} = I$:

$$\text{trace}(R) = \text{trace}(R\hat{V}^T\hat{V}) = \text{trace}\left(\hat{V}R\hat{V}^T\right) = \text{trace}(Y) = n + 1.$$
Set Constraints

**Cone constraints**

\[ R := \left\{ R \in \mathbb{S}^{(n-1)^2+1} : R \succeq 0, \text{trace}(R) = n + 1 \right\}, \]

**Polyhedral constraints**

\[ Y : = \left\{ Y \in \mathbb{S}^{n^2+1} : G_{\text{J}}(Y) = u_0, 0 \leq Y \leq 1, \right. \]
\[ \left. b^\circ \text{diag}(Y) = I, o^\circ \text{diag}(Y) = I, \text{arrow}(Y) = 0 \right\} \]
Main Model

(Split) Model

\[ p_{DNN}^* := \min_{R, Y} \langle L_Q, Y \rangle \]

\text{(DNN)}

\text{s.t.} \quad Y = \hat{V}R\hat{V}^T

R \in \mathcal{R}

Y \in \mathcal{Y}.

Recovering original \( X \) from \( \mathcal{Y} \) with redundant constraints

\( Y \in \mathcal{Y} \implies X = \text{Mat}(\text{diagonal}(Y)) = \text{Mat}(\text{Row1}(Y)) \) satisfy

\( Xe = X^Te = e; \) (so doubly stochastic)
Optimality Conditions

Lagrangian function, dual variable $Z$

$$\mathcal{L}(R, Y, Z) = \langle L_Q, Y \rangle + \langle Z, Y - \hat{V}R\hat{V}^T \rangle.$$ 

Strictly feasibility/onto $\hat{R} \succ 0$, $\hat{Y} = \hat{V}\hat{R}\hat{V}$ holds

first order optimality conditions:

1. $0 \in -\hat{V}^T Z \hat{V} + \mathcal{N}_R(R)$, (dual $R$ feasibility) (1a)
2. $0 \in L_Q + Z + \mathcal{N}_Y(Y)$, (dual $Y$ feasibility) (1b)
3. $Y = \hat{V}R\hat{V}^T$, $R \in \mathcal{R}$, $Y \in \mathcal{Y}$, (primal feasibility) (1c)

where $\mathcal{N}_R(R)$ (resp. $\mathcal{N}_Y(Y)$) is normal cone to $\mathcal{R}$ (resp. $\mathcal{Y}$) at $R$ (resp. $Y$).
For subproblems

The primal-dual $R, Y, Z$ are optimal if, and only if, normal cone conditions hold if, and only if,

$$R = \mathcal{P}_R(R + \hat{V}^T Z \hat{V}) \quad (2a)$$
$$Y = \mathcal{P}_Y(Y - L_Q - Z) \quad (2b)$$
$$Y = \hat{V} R \hat{V}^T. \quad (2c)$$
Let

$$Z_A := \left\{ Z \in S^{n^2+1} : (Z + L_Q)_{ij} = 0, \forall i, j \text{ (in arrow positions),} \right.$$ 
$$\text{and } \forall ij \in J_R \text{ (redundant gangster positions)} \right\}. $$
(restricted contractive Peaceman-Rachford splitting; redundant constraints in subproblems; modified dual variable)

**Augmented Lagrangian function**

\[
\mathcal{L}_A(R, Y, Z) = \langle L_Q, Y \rangle + \langle Z, Y - \hat{V}R\hat{V}^T \rangle + \frac{\beta}{2} \left\| Y - \hat{V}R\hat{V}^T \right\|_F^2,
\]

where \( \beta \) is a positive penalty parameter.

**Dual variable** \( Z \)

\[ Z_0 := \{ Z \in S^{n^2+1} : Z_{i,i} = 0, \ Z_{0,i} = Z_{i,0} = 0, \ i = 1, \ldots, n^2 \} \]

\( P_{Z_0} \) projection onto \( Z_0 \)
The Algorithm

**PRSM for DNN**

- **Initialize:** $\mathcal{L}_A$ is augmented Lagrangian; $\gamma \in (0, 1)$ is under-relaxation parameter; $\beta \in (0, \infty)$ is penalty parameter; $\mathcal{R}, \mathcal{Y}$ are subproblem sets; $Y^0$; and $Z^0 \in \mathcal{Z}_A$;

**WHILE** tolerances not met **DO**

- $R^{k+1} = \arg\min_{R \in \mathcal{R}} \mathcal{L}_A(R, Y^k, Z^k)$
- $Z^{k+\frac{1}{2}} = Z^k + \gamma\beta \cdot \mathcal{P}_{Z_0} \left( Y^k - \hat{V} R^{k+1} \hat{V}^T \right)$
- $Y^{k+1} = \arg\min_{Y \in \mathcal{Y}} \mathcal{L}_A(R^{k+1}, Y, Z^{k+\frac{1}{2}})$
- $Z^{k+1} = Z^{k+\frac{1}{2}} + \gamma\beta \cdot \mathcal{P}_{Z_0} \left( Y^{k+1} - \hat{V} R^{k+1} \hat{V}^T \right)$

**ENDWHILE**
### Algorithm outline/remarks

- alternate minimization of variables $R$ and $Y$ interlaced by the dual variable $Z$ update;
- $R$-update and the $Y$-update in are well-defined subproblems with unique solutions;
- many of the constraints are redundant in the SDP part but not within the subproblems; this improves rate of convergence and quality of $Y$ when stopping early.
- modified dual update $Z$ both after $R$-update and $Y$-update.

### Theorem

Let $\{R^k\}, \{Y^k\}, \{Z^k\}$ be the sequences generated by the algorithm. Then the sequence $\{(R^k, Y^k)\}$ converges to a primal optimal pair $(R^*, Y^*)$, and $\{Z^k\}$ converges to an optimal dual solution $Z^* \in Z_A$. 
R-subproblem

\[ \mathcal{R} := \left\{ R \in \mathbb{S}_+^{(n-1)^2+1} : \text{trace}(R) = n + 1 \right\}. \]

- \( \mathcal{P}_\mathcal{R}(W) \) projection of \( W \) onto \( \mathcal{R} \)
- completing the square at current \( Y^k, Z^k \): the \( R \)-subproblem can be explicitly solved by the projection operator \( \mathcal{P}_\mathcal{R} \) as follows:

\[
R^{k+1} = \arg\min_{R \in \mathcal{R}} -\langle Z^k, \hat{V} R \hat{V}^T \rangle + \frac{\beta}{2} \left\| Y^k - \hat{V} R \hat{V}^T \right\|_F^2
\]
\[
= \arg\min_{R \in \mathcal{R}} \frac{\beta}{2} \left\| Y^k - \hat{V} R \hat{V}^T + \frac{1}{\beta} Z^k \right\|_F^2
\]
\[
= \arg\min_{R \in \mathcal{R}} \frac{\beta}{2} \left\| R - \hat{V}^T (Y^k + \frac{1}{\beta} Z^k) \hat{V} \right\|_F^2
\]
\[
= \mathcal{P}_\mathcal{R}(\hat{V}^T (Y^k + \frac{1}{\beta} Z^k) \hat{V})
\]

- Eigendecomposition and projection onto simplex.
**Y-Subproblem**

\[ Y := \{ Y \in S^{n^2+1} : g_\mathcal{Y}(Y) = u_0, \ 0 \leq Y \leq 1, \ b^\mathcal{O}\text{diag}(Y) = I, \ o^\mathcal{O}\text{diag}(Y) = l, \ \text{arrow}(Y) = 0 \} \]

- \( \mathcal{P}_\mathcal{Y}(W) \) projection of \( W \) onto \( \mathcal{Y} \)
- completing the square at current \( R^{k+1}, Z^{k+\frac{1}{2}} \): the \( Y \)-subproblem can be explicitly solved by the projection operator \( \mathcal{P}_\mathcal{Y} \) as follows:

\[
Y^{k+1} = \arg\min_{Y \in \mathcal{Y}} \langle L_Q, Y \rangle + \langle Z^{k+\frac{1}{2}}, Y - \hat{V} R^{k+1} \hat{V}^T \rangle + \frac{\beta}{2} \left\| Y - \hat{V} R^{k+1} \hat{V}^T \right\|^2_F
\]

\[
= \arg\min_{Y \in \mathcal{Y}} \frac{\beta}{2} \left\| Y - \left( \hat{V} R^{k+1} \hat{V}^T - \frac{1}{\beta} (L_Q + Z^{k+\frac{1}{2}}) \right) \right\|^2_F
\]

\[
= \mathcal{P}_\mathcal{Y} \left( \hat{V} R^{k+1} \hat{V}^T - \frac{1}{\beta} \left( L_Q + Z^{k+\frac{1}{2}} \right) \right)
\]

- shooting holes; rounding to 0, 1
Provable Lower Bounds from Approximate Solutions

\[ g(Z) := \min_{Y \in \mathcal{Y}} \langle L_Q + Z, Y \rangle - (n + 1)\lambda_{\text{max}}(\hat{V}^T Z \hat{V}) \]

where \( \lambda_{\text{max}}(\hat{V}^T Z \hat{V}) \) denotes largest eigenvalue of \( \hat{V}^T Z \hat{V} \).

**Theorem**

\( d^*_Z := \max_Z g(Z) \) is a concave maximization problem. Furthermore, strong duality holds with main DNN problem:

\[ p^*_\text{DNN} = d^*_Z, \text{ and } d^*_Z \text{ is attained.} \]

\( Z \) dual feasible \( \implies g(Z) \) is a provable lower bound
Seemingly quadratic nearest discrete problem is a simplest LP

Given $\bar{X} \in \mathbb{R}^{n \times n}$

$$X^* = \arg\min_{X \in \Pi} \frac{1}{2} \| X - \bar{X} \|^2_F = \arg\min_{X \in \Pi} - \langle \bar{X}, X \rangle = \arg\min_{X \in D} - \langle \bar{X}, X \rangle,$$

since Von Neumann-Birkoff Theorem implies extreme points of doubly stochastic $D$ are the permutation matrices $\Pi$; so can apply a simplex method or Hungarian method for assignment problem.

Upper bound

A feasible solution $X^* \in \Pi$ to the original QAP, gives a valid upper bound $\text{trace}(AX^*B(X^*)^T)$. 
## Randomized Upper Bound

### Previous Approaches using an approximate optimum $Y^{out}$

(Exploit Perron-Frobenius to conclude $\nu_1 \geq 0$.)

1. $\text{vec } (\bar{X}) \cong \text{col. 1}( Y^{out})$; find nearest $X^* \in \Pi$.
2. $Y^{out} = \sum_{i=1}^{r} \lambda_i \nu_i \nu_i^T$ spectral decomposition,
   $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$; wlog $\nu_i \in \mathbb{R}^{n^2}$; $\text{vec } (\bar{X}) = \lambda_1 \nu_1$; find nearest $X^* \in \Pi$.

### Goemans-Williams type approximation algorithm, [3]

- $\xi \in (0, 1)^r$; in decreasing order; perturb eigenvalues;
- $\text{vec } (\bar{X}) = \sum_{i=1}^{r} \xi_i \lambda_i \nu_i$; find nearest $X^* \in \Pi$.
- repeat $\max\{1, \min(3 \cdot \lceil \log(n) \rceil), \text{ubest} - \text{lbest}\}$ number of times; ‘ubest’ and ‘lbest’ best current upper and lower bounds.
Numerical Experiments with PRSM

- recent: relaxation methods [1, C-SDP], [2, F2-RLT2-DA] and [5, SDPNAL].

sizes $n$: small, medium, large

$$n \in \{10, \ldots, 20\}, \{21, \ldots, 40\}, \{41, \ldots, 64\}.$$  

$n = 64: t(n^2 + 1) = 8,394,753$ variables; nonnegativity cuts; SDP constraints
Conclusion

- We introduced a strengthened splitting method for solving the facially reduced DNN relaxation for the QAP.
- Our strengthened model and algorithm incorporates redundant constraints to the model that are not redundant in the subproblems; more specifically, the trace constraint in $R$-subproblem and projection onto doubly stochastic matrices in $Y$-subproblem.
- We exploit the structure of dual optimal multipliers and provide customized dual updates; leads to a new strategy for strengthening the provable lower bounds.
- codes can be downloaded with link
  https://www.math.uwaterloo.ca/~7Ehwolkowi/henry/reports/ADMMnPRSMcodes.zip.


ADMM for the SDP relaxation of the QAP. 

L. Yang, D. Sun, and K.-C. Toh. 
SDPNAL+: a majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. 

Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. 
Semidefinite programming relaxations for the quadratic assignment problem. 

Semidefinite Programming and Interior-point Approaches for Combinatorial Optimization Problems (Fields Institute, Toronto, ON, 1996).
Thanks for your attention!

A Restricted Dual Peaceman-Rachford Splitting Method for a Strengthened DNN Relaxation for QAP

Henry Wolkowicz
Dept. Comb. and Opt., University of Waterloo, Canada

Thur. May 19, 13:45-14:35 (BST) (08:45-09:35 EDT) 2022

At: Modern Techniques of Very Large Scale Optimization

joint with: Naomi Graham, Hao Hu, Jiyoung Im, Xinxin Li

2