Input:

```
Output:
x \in R^p V \in R^{k \times p} k < p
x \in R^{p} f: R^{p} \to R\hat{y} = w_{0} + w^{T} x + \sum_{i=1}^{p} \sum_{j=i+1}^{p} v_{i}^{T} v_{j} x_{i} x_{j}
\hat{y} = w_{0} + w^{T} x + x^{T} W x v_{i} R^{k} V V = (v_{1}, v_{2}, ..., v_{p}) \sum_{i=1}^{p} \sum_{j=i+1}^{p} v_{i}^{T} v_{j} x_{i} x_{j} = x^{T} V V^{T} x - x^{T} diag(V V^{T}) x = x^{T} (V V^{T} - diag(V V^{T}) x)
\pi_i = \langle W, x_i x_i^T \rangle
```

 $wWf_1(w)\nabla f_1(w) = -X'(y - diag(XZX') - Xw) + \alpha ww^+ = w - t\nabla f_1(w)wg(W) + h(W)g(W) = \sum_{i=1}^n 1/2(y_i - w^Tx_i - x_i^Tx_i)$

 $W = U\Sigma V^T \Sigma_{\beta t}$ g(W) $0 < \beta < 1$ $t = \beta t$ $WZWW - Z = 0 \|W\|_{tr} \|Z\|_{tr0}, w, W, Z, u) = f(w_0, w, W) + \lambda_1 \|Z\|_{tr} + \lambda_2 \|vec(W)\|_1 + \lambda_3 \|w\|_2^2 + \langle W - Z, u \rangle + m \|W - Z\|_F^2 u \in \mathbb{R}^{p \times p} m$ $w_0 = u + W - Z$ $w_0, w, W w_0, w W_0, w = \arg\min_{w_0, w} f(w_0, w, W) + \lambda_3 \|w\|_2^2$ $W = \arg\min_{w_0, w_0} f(w_0, w, W) + \lambda_2 \|vec(W)\|_1 + \frac{\rho}{2} \|W - Z + u\|_F^2 w_0, wWW$ $w_0^0, w^0, W^0, Z^0, u^0, \rho w_0, w^k, Wk = 1 > w_0^k, w^k, W^k Z^k = prox_{tr, \frac{\lambda_1}{\rho}} (W^k + u^{k-1})u^k = W^k - Z^k$

 $_cfmpresents the results in solving eq: criteria 1 with method proposed in sec: cfm. sec: result_scfmpresents the results in inverse in the results in the$ $_f$ igs shows the convergence rate of four data types. The order of the convergence rate is as follows: $SYMSPARSEASYM \approx$

 $||W - Z||_F^2$

 $p10, 30, 100, 500, 1000WW = V^TVW = U^TVWW$

 $\stackrel{w_0}{W}\stackrel{v}{W}\stackrel{v}{W}\stackrel{v}{W}\stackrel{v}{k}$ Proceedings of the 24th international conference on Machine learning Joint European Conference on Machine Learning and